

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 94.3%; Score 66; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred.No. 3.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
| | | | | | | | | |  
Db 1 CFQWQNRKRV 11

RESULT 2  
US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,948  
FILING DATE: APRIL 4, 1996  
CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995  
APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-Cl, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 94.3%; Score 66; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred.No. 3.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
| | | | | | | | | |  
Db 1 CFQWQNRKRV 11

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,380  
FILING DATE: April 4, 1996  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-Cl, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 94.3%; Score 66; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred.No. 3.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
| | | | | | | | | |  
Db 1 CFQWQNRKRV 11

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/488,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LF-Cl, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

Query Match 94.3%; Score 66; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRMKRV 11
Db 1 CFQWQRMKRV 11

RESULT 5
US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500kb
COMPUTER: IBM Compatible

```

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3
Query Match 94.3%; Score 66; DB 1; Length 20;

```

Best Local Similarity 100.0%; Pred. No. 3.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRKRV 11  
| | | | | | | | | |  
Db 2 CFQWQNRKRV 12

## RESULT 6

US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/891,174

FILING DATE: 29-MAY-1992

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/755,161

FILING DATE: 05-SEP-1991

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM:

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 94.3%; Score 66; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRKRV 11

| | | | | | | | | |  
Db 2 CFQWQNRKRV 12

## RESULT 7

US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,487  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A.  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: RJN-019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOPERRIN"  
US-08-204-487-1

Query Match 94.3%; Score 66; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMRKV 11  
Db 2 CFQWQNNMRKV 12

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"

US-08-256-771-24  
Query Match 94.3%; Score 66; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMRKV 11  
Db 2 CFQWQNNMRKV 12

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
US-08-256-771-25

Query Match 94.3%; Score 66; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMRKV 11  
Db 2 CFQWQNNMRKV 12

RESULT 10  
US-08-381-984-24



```

; Sequence 24, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"
; US-08-381-984-24
;
; Query Match 94.3%; Score 66; DB 1; Length 20;
; Best Local Similarity 100.0%; Pred. No. 3.7e-05;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 CFQWQNRMRKV 11
; Db 2 CFQWQNRMRKV 12
;
; RESULT 11
; US-08-381-984-25
; Sequence 25, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack

```

```

; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"
; US-08-381-984-25
;
; Query Match 94.3%; Score 66; DB 1; Length 20;
; Best Local Similarity 100.0%; Pred. No. 3.7e-05;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 CFQWQNRMRKV 11
; Db 2 CFQWQNRMRKV 12
;
; RESULT 12
; US-09-508-734-4
; Sequence 4, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; TITLE OF INVENTION: useful microorganism thereof
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 4

```

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match          94.3%; Score 66; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11
Db 2 CFQWQRMNRKV 12

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: useful microorganism thereof
; CURRENT FILING DATE: 2000-06-01
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          94.3%; Score 66; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11
Db 3 CFQWQRMNRKV 13

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8950
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match          94.3%; Score 66; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11
Db 4 CFQWQRMNRKV 14
```

## RESULT 15

```
US-07-891-174-10
; Sequence 10, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
```

```
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
; Query Match 94.3%; Score 66; DB 1; Length 25;
; Best Local Similarity 100.0%; Pred. No. 4.6e-05;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; QY 1 CFQWQRMNRKV 11
; Db 4 CFQWQRMNRKV 14
; Search completed: February 21, 2003, 07:50:35
; Job time : 8.7 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-79  
Perfect score: 70  
Sequence: 1 CFQWQRNMRKVA 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query *	Score	Match	Length	ID	Description
1	66	94.3	15	9	US-09-798-869-2	Sequence 2, Appli
2	66	94.3	25	9	US-09-798-869-20	Sequence 20, Appli
3	66	94.3	694	9	US-10-023-096-2	Sequence 2, Appli
4	58	82.9	15	9	US-09-798-869-6	Sequence 6, Appli
5	54	77.1	15	9	US-09-798-869-3	Sequence 3, Appli
6	54	77.1	25	9	US-09-798-869-23	Sequence 23, Appli
7	46	65.7	15	9	US-09-798-869-7	Sequence 7, Appli
8	45	64.3	15	9	US-09-798-869-4	Sequence 4, Appli
9	45	64.3	25	9	US-09-798-869-22	Sequence 22, Appli
10	42	60.0	15	9	US-09-798-869-8	Sequence 8, Appli
11	42	60.0	15	9	US-09-798-869-29	Sequence 29, Appli
12	42	60.0	15	9	US-09-798-869-30	Sequence 30, Appli
13	38	54.3	489	9	US-09-888-320-2	Sequence 2, Appli
14	37	52.9	21	10	US-09-864-761-47985	Sequence 47985, A
15	37	52.9	747	9	US-10-066-500-58	Sequence 58, Appli
16	37	52.9	747	9	US-10-002-796-58	Sequence 58, Appli
17	37	52.9	747	9	US-10-066-273-58	Sequence 58, Appli
18	37	52.9	747	9	US-10-066-494-58	Sequence 58, Appli
19	36	51.4	209	10	US-09-904-536-8	Sequence 8, Appli

20	36	51.4	209	10	US-09-904-536-9	Sequence 9, Appli
21	36	51.4	209	10	US-09-904-536-11	Sequence 11, Appli
22	36	51.4	209	10	US-09-904-536-12	Sequence 12, Appli
23	36	51.4	209	10	US-09-904-536-13	Sequence 13, Appli
24	36	51.4	209	10	US-09-904-536-14	Sequence 14, Appli
25	36	51.4	209	10	US-09-904-536-15	Sequence 15, Appli
26	36	51.4	209	10	US-09-904-536-16	Sequence 16, Appli
27	36	51.4	209	10	US-09-904-536-17	Sequence 17, Appli
28	36	51.4	209	10	US-09-904-536-18	Sequence 18, Appli
29	36	51.4	212	10	US-09-904-536-10	Sequence 10, Appli
30	36	51.4	235	9	US-10-095-449-6	Sequence 6, Appli
31	36	51.4	235	10	US-09-448-378-1	Sequence 1, Appli
32	36	51.4	235	10	US-09-983-806-6	Sequence 6, Appli
33	36	51.4	235	10	US-09-904-536-1	Sequence 1, Appli
34	36	51.4	338	9	US-09-978-295A-119	Sequence 119, App
35	36	51.4	338	9	US-09-978-697-119	Sequence 119, App
36	36	51.4	338	9	US-09-978-192A-119	Sequence 119, App
37	36	51.4	338	9	US-09-999-832A-119	Sequence 119, App
38	36	51.4	338	9	US-09-978-189-119	Sequence 119, App
39	36	51.4	553	9	US-09-796-753-14	Sequence 14, Appli
40	36	51.4	553	10	US-09-981-649A-6	Sequence 6, Appli
41	36	51.4	553	10	US-09-981-649A-24	Sequence 24, Appli
42	36	51.4	554	10	US-09-981-649A-30	Sequence 30, Appli
43	36	51.4	554	10	US-09-981-649A-32	Sequence 32, Appli
44	36	51.4	559	10	US-09-981-649A-28	Sequence 28, Appli
45	36	51.4	607	9	US-09-881-579-10	Sequence 10, Appli

## ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication NO. US2003002821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 94.3%; Score 66; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRNMRKV 11  
Db 3 CFQWQRNMRKV 13

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US2003002821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 94.3%; Score 66; DB 9; Length 25;  
Best Local Similarity 100.0%; Pred. No. 5.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11  
| | | | | | | | | |  
Db 3 CFQWQRMNRKV 13

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,585  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 94.3%; Score 66; DB 9; Length 694;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11  
| | | | | | | | | |  
Db 22 CFQWQRMNRKV 32

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 82.9%; Score 58; DB 9; Length 15;  
Best Local Similarity 90.9%; Pred. No. 0.00072;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKV 11  
| | | | | | | | | |  
Db 3 CFQWQRMNRKV 13

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 77.1%; Score 54; DB 9; Length 15;  
Best Local Similarity 72.7%; Pred. No. 0.0034;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|||||:  
Db 3 CYQWQRMKRL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 77.1%; Score 54; DB 9; Length 25;  
Best Local Similarity 72.7%; Pred. No. 0.0054; 1; Indels 0; Gaps 0;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|||||:  
Db 3 CYQWQRMKRL 13

## RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 65.7%; Score 46; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.073; 2; Indels 0; Gaps 0;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|||||:  
Db 3 CYQWQRMKRL 13

## RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 64.3%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.11; 3; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|||||:  
Db 3 CLRWQRMKRV 13

## RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 64.3%; Score 45; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.17; 3; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|||||:  
Db 3 CLRWQRMKRV 13

```
RESULT 10
US-09-798-869-8
; Sequence 8, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine
; OTHER INFORMATION: sequence)
US-09-798-869-8

Query Match      60.0%; Score 42; DB 9; Length 15;
Best Local Similarity 63.6%; Pred. No. 0.34;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRMKV 11
DB 3 CLRQWEMKVL 13

RESULT 11
US-09-798-869-29
; Sequence 29, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 15
; TYPE: PRT
; ORGANISM: BOVINE
US-09-798-869-29

Query Match      60.0%; Score 42; DB 9; Length 15;
Best Local Similarity 54.5%; Pred. No. 0.34;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNRMKV 11
DB 3 CLRQWEMKVL 13

RESULT 12
US-09-798-869-30
; Sequence 30, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 15
; TYPE: PRT
; ORGANISM: BOVINE
US-09-798-869-30

Query Match      60.0%; Score 42; DB 9; Length 15;
Best Local Similarity 54.5%; Pred. No. 0.34;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNRMKV 11
DB 3 CLRQWEMKVL 13

RESULT 13
US-09-888-320-2
; Sequence 2, Application US/09888320
; Publication No. US20030013090A1
; GENERAL INFORMATION:
; APPLICANT: Barry III, Clifton E.
; APPLICANT: DeBarber, Andrea E.
; APPLICANT: Mdulali, Khisimuzi
; APPLICANT: Bekker, Linda-Gail
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis
; FILE REFERENCE: 015280-413100US
; CURRENT APPLICATION NUMBER: US/09/888,320
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/214,187
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: wild-type EtAa monoxygenase (Rv3854c, EthA)
US-09-888-320-2

Query Match      54.3%; Score 38; DB 9; Length 489;
Best Local Similarity 54.5%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRMKV 11
DB 253 CQKPRPRMKRM 263

RESULT 14
US-09-864-761-47985
; Sequence 47985, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aetmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47985
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096701.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EST_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06
; US-09-864-761-47985

Query Match 52.9%; Score 37; DB 10; Length 21;
Best Local Similarity 83.3%; Pred. No. 3.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWR 6
Db 16 CFQWR 21

RESULT 15
US-10-066-500-58
; Sequence 58, Application US/10066500
; Patent No. US2002017165A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Botstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Eaton
; APPLICANT: Napoleon Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3130RIC7
; CURRENT APPLICATION NUMBER: US/10/066,500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
```



; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100858  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101922  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/106032  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/109304  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/125778  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/139695  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/145070  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/149396  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 08/918874  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 08/933821  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 08/960507  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 09/114844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: 09/136801  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136804  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136828  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/158342  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: 09/180997  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 09/202088  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: 09/254311  
; PRIOR FILING DATE: 1999-03-03  
; PRIOR APPLICATION NUMBER: 09/254460  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: 09/254465  
; PRIOR FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: 09/284663  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 09/332928  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/332929  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333075  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333077  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380139  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/403296  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/403297  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/423741  
; PRIOR FILING DATE: 1999-11-10  
; PRIOR APPLICATION NUMBER: 09/423844  
; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 09/522342  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 09/548815  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: 09/664610  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/665350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/709238  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/767609  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 09/802706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 09/808689  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/866028  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 09/870574  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 09/872035  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/886342  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: PCT/US98/14552  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US98/18824  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/US98/19093  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: PCT/US98/19330  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: PCT/US98/19437  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/24855  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: PCT/US98/25108  
; PRIOR FILING DATE: 1998-12-01  
; PRIOR APPLICATION NUMBER: PCT/US98/25190  
; PRIOR FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: PCT/US99/20111  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 52.9%; Score 37; DB 9; Length 747;  
Best Local Similarity 45.5%; Pred. No. 91;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRKVV 11

Db 311 CVRWQINSRRI 321

Search completed: February 21, 2003, 08:08:07  
Job time : 11.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds

(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107b-79

Perfect score: 70

Sequence: 1 CFQWRNRKVA 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	66	94.3	711	1 TFHUL	lactotransferrin p
2	54	77.1	708	2 JC2323	lactoferrin - goat
3	51	72.9	33	2 S52107	lactoferrin - shee
4	45	64.3	707	1 A28438	lactoferrin precu
5	43	61.4	4568	2 T08030	dynein beta heavy
6	42	60.0	298	2 AD2346	hypothetical prote
7	41	58.6	500	2 S42867	protein kinase (EC
8	39	55.7	121	2 AH3147	hypothetical prote
9	39	55.7	435	2 S52784	ornithine decarbox
10	39	55.7	932	2 T28820	hypothetical prote
11	38	54.3	205	2 E90094	26S proteasome SU
12	38	54.3	282	2 F90580	hypothetical prote
13	38	54.3	397	2 T35361	hypothetical prote
14	38	54.3	464	2 AI2343	hypothetical prote
15	38	54.3	489	2 C70655	probable monooxyge
16	38	54.3	515	2 T00510	probable cytochrom
17	38	54.3	543	2 T00513	cytochrome P450 ho
18	38	54.3	966	1 P1RVBB	RNA la protein - b
19	37	52.9	303	2 T90848	probable oxidoredu
20	37	52.9	361	2 T29571	hypothetical prote
21	37	52.9	511	2 AB0858	hypothetical prote
22	37	52.9	531	2 A84471	En/spm-like transp
23	37	52.9	566	2 S75233	ABC transporter sl
24	37	52.9	570	2 T45261	hypothetical prote
25	37	52.9	829	2 C82361	GDGEF family prote
26	37	52.9	1051	2 T48933	WD repeat domain p
27	36	51.4	119	2 AE0040	conserved hypothet
28	36	51.4	124	2 C96582	F1511.22 (imported
29	36	51.4	211	2 D82109	outer membrane lip

## ALIGNMENTS

### RESULT 1

#### TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence\_revision 21-Nov-1997 #text\_change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R:Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R:Rev. M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A>Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148,'T',150-422,'C',424-711 <REV>

A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A>Title: Differential molecular mechanism of the estrogen action that regulates lactoferrin

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A:Experimental source: placenta

A>Note: sequence extracted from NCBI backbone (NCBIP:122202)

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A>Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A>Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <ST1>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-28,'X',30-31 <ST2>

R:Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A:Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A:Reference number: S07160; MUID:89001031; PMID:3477300  
 A:Accession: S07160  
 A:Molecule type: mRNA  
 A:Residues: 436-487, 'A' 489-711 <RAD>  
 A:CROSS-references: EMBL:M18642; NID:gl86815; PIDN:AAA86665.1; PID:G386855  
 R:Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A:Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A:Reference number: A61169; MUID:91235214; PMID:1674448  
 A:Accession: A61169  
 A:Molecule type: mRNA  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 3-701, 'SWKPVN' <PAN>  
 A:Experimental source: normal breast tissue  
 R:Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A:Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A:Reference number: A31000; MUID:85076667; PMID:6510420  
 A:Accession: A31000  
 A:Molecule type: protein  
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A:Note: this is the final paper in a series  
 R:Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A:Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A:Reference number: S74119; MUID:97054624; PMID:8898921  
 A:Accession: S74119  
 A:Molecule type: protein  
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A:Experimental source: neutrophil granulocytes  
 C:Genetics:  
 A:Gene: GDB:LTF  
 A:CROSS-references: GDB:119368; OMIM:150210  
 A:Map position: 3q21-3q23  
 C:Superfamily: transferrin; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-711/Product: lactotransferrin #status experimental <MAT>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-637, 595-609/Disulfide bonds: #status e  
 F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-652/Disulfide bonds: #stat

Query Match 94.3%; Score 66; DB 1; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
 |||||  
 DB 39 CFQWQNRKRV 49

RESULT 2  
 JC2323  
 lactoferrin - goat  
 C:Species: Capra aegagrus hircus (domestic goat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C:Accession: JC2323  
 R:Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
 Biochim. Biophys. Res Commun. 203, 1324-1332, 1994  
 A:Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A:Reference number: JC2323; MUID:94380047; PMID:8093048  
 A:Accession: JC2323  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <LEP>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:359-696/Domain: transferrin repeat homology <TRH2>  
 F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 77.1%; Score 54; DB 2; Length 708;  
 Best Local Similarity 72.7%; Pred. No. 0.12;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
 |||||  
 DB 38 CFQWQNRKRV 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C:Accession: S52107  
 R:Qian, Z.Y.; Jolles, P.; Migliore-Samouri, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A:Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet a  
 A:Reference number: S52107; MUID:95127729; PMID:7827104  
 A:Accession: S52107  
 A:Molecule type: protein  
 A:Residues: 1-33 <QIA>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication

Query Match 72.9%; Score 51; DB 2; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.019;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
 |||||  
 DB 19 CYQWQNRKRV 29

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N:Alternate names: lactotransferrin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A28438; A41205  
 R:Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A:Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secre  
 A:Reference number: A92596; MUID:87280033; PMID:3611056  
 A:Accession: A28438  
 A:Molecule type: mRNA  
 A:Residues: 3-707 <PEN>  
 A:CROSS-references: EMBL:J03298  
 R:Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A:Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A:Reference number: A41205; MUID:92042099; PMID:1939212  
 A:Accession: A41205  
 A:Molecule type: DNA  
 A:Residues: 1-15 <LIU>  
 A:CROSS-references: GB:M74778  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-707/Product: lactotransferrin #status predicted <MAT>  
 F:358-695/Domain: transferrin repeat homology <TRH2>  
 F:7494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.3%; Score 45; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 4.6;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11  
 |||||  
 DB 37 CLRWQNRKRV 47

```

RESULT 5
T08030
dynein beta heavy chain - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 02-Feb-2001
C:Accession: T08030
R:Michell, D.R.; Brown, K.S.
J. Cell Sci. 107, 635-644, 1994
A:Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.
A:Reference number: Z16302; MUID:94274778; PMID:8006077
A:Accession: T08030
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4568 <MIT>
A:Cross-references: EMBL:U02963; NID:G409965; PIDN:AAA19956.1; PID:G514215
A:Experimental source: strain 2lgr
C:Genetics:
A:Gene: ODA4
A:Map position: IX
A:Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;
3334/3; 3686/3; 3882/3; 4240/3
C:Superfamily: dynein heavy chain, ciliary
C:Keywords: nucleotide binding; P-loop
F:1919-1926/Region: nucleotide-binding motif A (P-loop)
F:2202-2209/Region: nucleotide-binding motif A (P-loop)
F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 61.4%; Score 43; DB 2; Length 4568;
Best Local Similarity 54.5%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRKRV 11
|||||:|
Db 1852 CFQWQSLRYI 1862

RESULT 6
AD2346
hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AD2346
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076022.1; PID:G17133459; GSPDB:GNC0179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4323

Query Match 60.0%; Score 42; DB 2; Length 298;
Best Local Similarity 77.8%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWRNRK 10
|||||
Db 163 FHWQNRK 171

RESULT 7
S42867
protein kinase (EC 2.7.1.1) - spinach
C:Species: Spinacia oleracea (spinach)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 19-Jul-2002
C:Accession: S42867
R:Baur, B.; Winter, K.; Fischer, K.; Dietz, K.
submitted to the EMBL Data Library, March 1994
A:Description: Molecular cloning and characterization of several protein kinases from pl
A:Reference number: S42864
A:Accession: S42867
A:Molecule type: mRNA
A:Residues: 1-500 <BAU>
A:Cross-references: EMBL:Z30330; NID:G457708; PIDN:CAA82991.1; PID:G457709
C:Superfamily: Arabidopsis thaliana protein kinase T518.9; protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase
F:71-379/Domain: protein kinase homology <KIN>
F:79-87/Region: protein kinase ATP-binding motif

Query Match 58.6%; Score 41; DB 2; Length 500;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQNRKVA 12
|||||:|
Db 268 WQNRKLA 276

RESULT 8
AH3147
hypothetical protein Atu4804 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AH3147
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AH3147
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <KUR>
A:Cross-references: PIDN:AAL45598.1; PID:G17743317; GSPDB:GNC00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4804
A:Map position: linear chromosome

Query Match 55.7%; Score 39; DB 2; Length 121;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQNRKVA 12
|||||:|
Db 14 CLAWQNRNRVS 25

RESULT 9
S52784
ornithine decarboxylase (EC 4.1.1.17) - Panagrellus redivivus
C:Species: Panagrellus redivivus
C>Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 08-Sep-2000
C:Accession: S55347; S52784
R:yon Besser, H.; Nienann, G.; Domdey, B.; Walter, R.D.
Biochem. J. 308, 635-640, 1995
A:Title: Molecular cloning and characterization of ornithine decarboxylase cDNA of the n
A:Reference number: S55347; MUID:95290001; PMID:7772052
A:Accession: S55347
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-435 <VON>
A:Cross-references: EMBL:X82199; NID:G758641; PIDN:CAA57683.1; PID:G758642
C:Superfamily: ornithine decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis; l

```

F7/76/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 55.7%; Score 39; DB 2; Length 435;  
Best Local Similarity 70.0%; Pred. No. 33;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQNRKVA 12  
||| | | |  
DB 61 QWQNRKVA 70

## RESULT 10

T28820

hypothetical protein F07C3.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T28820

R:Favella, A.; Gattung, S.

submitted to the EMBL Data Library, March 1996

A:Description: The sequence of *C. elegans* cosmid F07C3.

A:Reference number: Z20528

A:Accession: T28820

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-932 <FAV>

A:Cross-references: EMBL:U50308; PIDN:RAC48001.1; GSPDB:GN000023; CESP:F07C3.1

A:Experimental source: strain Bristol N2; clone F07C3

C:Genetics:

A:Gene: CESP:F07C3.1

A:Map position: 5

A:Introns: 14/2; 55/3; 86/2; 159/3; 220/2; 268/3; 327/1; 363/3; 402/3; 463/2; 559/3; 599

## Query Match

55.7%; Score 39; DB 2; Length 932;

Best Local Similarity 70.0%; Pred. No. 70;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOWQNRKXV 11

||||| |

DB 579 FQWQNRKLV 588

## RESULT 11

E90094

26S proteasome SU B5 [imported] - *Guillardia theta* nucleomorph

C:Species: nucleomorph *Guillardia theta*

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001

C:Accession: E90094

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: E90094

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-205 <DOU>

A:Cross-references: GB:AF165818; NID:gl3794510; PIDN:AAK39885.1; GSPDB:GN00150

C:Genetics:

A:Gene: prsB5

A:Map position: 1

A:Genome: nucleomorph

C:Keywords: nucleomorph

## Query Match

54.3%; Score 38; DB 2; Length 205;

Best Local Similarity 62.5%; Pred. No. 23;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWQNRN 8

||||| |

DB 63 CFQWERNL 70

## RESULT 12

F90580

hypothetical protein MYPV\_5500 [imported] - *Mycoplasma pulmonis* (strain UAB CTIP)

C:Species: *Mycoplasma pulmonis*

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: F90580

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pul*

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: F90580

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-282 <KUR>

A:Cross-references: GB:AL445566; PID:gl4089965; PIDN:CAC13723.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV\_5500

A:Genetic code: SGC3

## Query Match

54.3%; Score 38; DB 2; Length 282;

Best Local Similarity 50.0%; Pred. No. 32;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOWQNRKXV 11

||||| |

DB 20 FAWQNRKXV 29

## RESULT 13

T35361

hypothetical protein SC66T3.04 - *Streptomyces coelicolor*

C:Species: *Streptomyces coelicolor*

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C:Accession: T35361

R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21576

A:Accession: T35361

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-397 <MUR>

A:Cross-references: EMBL:AL079348; PIDN:CAB45460.1; GSPDB:GN000070; SCOEDB:SC66T3.04

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC66T3.04

## Query Match

54.3%; Score 38; DB 2; Length 397;

Best Local Similarity 66.7%; Pred. No. 45;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQNRKXV 12

||||| |

DB 206 WQNRKXV 214

## RESULT 14

A12343

hypothetical protein all4304 [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.

A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: A12343

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saamamoto, S.; Watanabe, A.; Iriguchi;

Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: A12343

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-464 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076003.1; PID:gl7133440; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:  
A:Gene: all4304

Query Match 54.3%; Score 38; DB 2; Length 464;  
Best Local Similarity 54.5%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWQNNRKVA 12  
||| | : :  
Db 378 FQWQNNKRAA 388

## RESULT 15

C70655  
probable monooxygenase - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: C70655  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID:98295987; PMID:9634230  
A: Accession: C70655  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-489 <COL>  
A: Cross-references: GB:Z83864; GB:AL123456; NID:g3261687; PIDN:CAB06212.1; PID:e301250;  
A: Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: Rv3854c

Query Match 54.3%; Score 38; DB 2; Length 489;  
Best Local Similarity 54.5%; Pred. No. 56;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNNRKV 11  
| : | | | :  
Db 253 CQKPRRRMKX 263

Search completed: February 21, 2003, 07:47:51  
Job time : 10.65 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 seconds  
(without alignments)  
108.139 Million cell updates/sec

Title: US-09-743-107B-79  
Perfect score: 70  
Sequence: 1 CFQWRNRKVA 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	94.3	711	1	TRFL_HUMAN
2	54	77.1	708	1	TRFL_CAPHI
3	49	70.0	708	1	TRFL_CAMDR
4	45	64.3	707	1	TRFL_MOUSE
5	43	61.4	4568	1	DVH8_CHLRE
6	41	58.6	695	1	TRFL_HORSE
7	39	55.7	435	1	DCOR_PANRE
8	38	54.3	146	1	RPOB_LTBAF
9	38	54.3	966	1	VIA_BBMV
10	37	52.9	435	1	DHOM_METGL
11	37	52.9	455	1	YKYL_CAEEL
12	37	52.9	566	1	YJ19_SNNY3
13	36	51.4	211	1	LOUB_VIBCH
14	36	51.4	214	1	VIF_SIVS4
15	36	51.4	235	1	FL3L_HUMAN
16	36	51.4	306	1	BUB2_YEAST
17	36	51.4	485	1	GLCA_BACST
18	36	51.4	1135	1	PHYC_SORBI
19	36	51.4	2671	1	IP3T_HUMAN
20	35	50.0	85	1	PMRD_SALTY
21	35	50.0	275	1	IL2A_BOVIN
22	35	50.0	275	1	IL2A_SHEEP
23	35	50.0	365	1	LA34_HUMAN
24	35	50.0	428	1	SYH_CHLMU
25	35	50.0	502	1	C911_AKATH
26	35	50.0	663	1	PD11_HUMAN
27	35	50.0	708	1	TRFL_BUBBU
28	35	50.0	728	1	KDGI_ARATH
29	35	50.0	765	1	Y008_HUMAN
30	35	50.0	783	1	YNR2_CAEEL
31	35	50.0	961	1	VIA_BMV
32	35	50.0	962	1	YBX7_SCHPO
33	35	50.0	1241	1	NPHN_HUMAN

34	34	48.6	152	1	YB83_METUA	Q58878 methanococ
35	34	48.6	192	1	RL24_SCHPO	Q10353 schizosacch
36	34	48.6	215	1	VIF_HV2SB	P12452 human immun
37	34	48.6	215	1	VIF_HV2ST	P20878 human immun
38	34	48.6	246	1	Y495_SYNY3	Q55185 synecocyst
39	34	48.6	275	1	VAI6_VACCV	P16710 vaccinia vi
40	34	48.6	275	1	VNS2_DSDNV	O71154 diatraea sa
41	34	48.6	292	1	NLA_DROME	Q9xz18 drosophila
42	34	48.6	316	1	NORC_CHLTR	O84281 chlamydia t
43	34	48.6	329	1	CATK_RAT	O35186 rattus norv
44	34	48.6	355	1	MORG_NEIMA	Q9Jsz7 neisseria m
45	34	48.6	355	1	MORG_NEIME	Q9K0Y2 neisseria m

## ALIGNMENTS

RESULT 1  
TRFL\_HUMAN STANDARD; PRT; 711 AA.  
AC P02788; Q16780; Q16785; Q16789; Q00756; Q9H123; Q96K24;  
AC Q96K25;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;  
DE Lactoferrin B; Lactoferrin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cho Y.Y.;  
RT Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Connely O.M.;  
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Liang O., Jimenez-Flores R., Richardson T.;  
RT "Molecular cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences.";  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RT "cDNA cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90326549; PubMed=2374734;  
 RA Powell M.J., Ogden J.E.;  
 RT "Nucleotide sequence of human lactoferrin cDNA.";  
 RL Nucleic Acids Res. 18:4013-4013(1990).  
 RN [9]  
 RP SEQUENCE OF 20-711.  
 RX MEDLINE=85076667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legrand D., Spik G., Montreuil J., Jolles P.;  
 RT "Human lactotransferrin: amino acid sequence and structural  
 RT comparisons with other transferrins.";  
 RL Eur. J. Biochem. 145:659-666(1984).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RX MEDLINE=82046817; PubMed=6794640;  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactotransferrin sequence. Study and  
 RT alignment of the cyanogen bromide fragments and characterization of  
 RT N- and C-terminal domains.";  
 RL Biochim. Biophys. Acta 670:243-254(1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RX MEDLINE=82262043; PubMed=7049727;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 RT lactotransferrin.";  
 RL FEBS Lett. 142:107-110(1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RX MEDLINE=88001031; PubMed=3477300;  
 RA Rado T.A., Wei X., Benz E.J. Jr.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 RT expression of mRNA during normal and leukemic myelopoiesis.";  
 RL Blood 70:989-993(1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
 RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
 RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
 RA Sagripanti J.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=90064528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 RT and refinement at 2.8-A resolution.";  
 RL J. Mol. Biol. 209:711-734(1989).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA Haridas M., Anderson B.F., Baker E.N.;  
 RT "Structure of human diferric lactoferrin refined at 2.2-A  
 RT resolution.";  
 RL Acta Crystallogr. D 51:629-646(1995).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
 RX MEDLINE=97156796; PubMed=9003186;  
 RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
 RA Baker E.N.;  
 RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
 RT binding properties and crystal structure of the histidine-  
 RT 253-->methionine mutant.";  
 RL Biochemistry 36:341-346(1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=99190892; PubMed=10089347;  
 RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
 RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
 RT awamori.";  
 RL Acta Crystallogr. D 55:403-407(1999).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=99192677; PubMed=10089508;  
 RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
 RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
 RT and analysis of ligand-induced conformational change.";  
 RL Acta Crystallogr. D 54:1319-1335(1998).  
 RN [19]  
 RP CHARACTERIZATION OF LACTOFERROXINS.  
 RX MEDLINE=91166929; PubMed=1369293;  
 RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
 RT "Isolation and characterization of opioid antagonist peptides derived  
 RT from human lactoferrin.";  
 RL Agric. Biol. Chem. 54:1803-1810(1990).  
 RN [20]  
 RP VARIANTS THR-30 AND ARG-48.  
 RX PubMed=9873069;  
 RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
 RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
 RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
 RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
 RA Hejtmancik J.F., Teng C.T.;  
 RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
 RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
 RL Mol. Vision 4:31-32(1998).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
 CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
 CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
 CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



Query Match 94.3%; Score 66; DB 1; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 0.00027;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
 |||||  
 Db 39 CFQWQRMKRV 49

## RESULT 2

TRFL\_CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA MEDLINE=94380047; PubMed=8093048;  
 RA Le Provost F., Nocard M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 RT relevant locus to bovine U12 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U53857; AAA37958.1; -;  
 DR EMBL; X78902; CAA55517.1; -;  
 DR HSP; O77698; 1CE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin.  
 DR PRINTS; PR00422; Transferrin; 2.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.

FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.  
 FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 56 56 I -> V (IN REF. 2).  
 FT CONFLICT 88 88 L -> R (IN REF. 2).  
 FT CONFLICT 124 124 Q -> K (IN REF. 2).  
 FT CONFLICT 154 154 F -> P (IN REF. 2).  
 FT CONFLICT 304 304 S -> R (IN REF. 2).  
 FT CONFLICT 414 414 D -> G (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;  
 Query Match 77.1%; Score 54; DB 1; Length 708;  
 Best Local Similarity 72.7%; Pred. No. 0.037;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQRMKRV 11  
 |||||  
 Db 38 CFQWQRMKRV 48  
 RESULT 3  
 TRFL\_CAMDR STANDARD; PRT; 708 AA.  
 ID TRFL\_CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9MZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) Lactoferrin.";  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary Gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.



```

FT REPEAT 358 707
FT DISULFID 27 63
FT DISULFID 37 54
FT DISULFID 133 216
FT DISULFID 175 191
FT DISULFID 188 199
FT DISULFID 249 263
FT DISULFID 366 398
FT DISULFID 376 389
FT DISULFID 423 702
FT DISULFID 443 665
FT DISULFID 475 550
FT DISULFID 499 693
FT DISULFID 509 523
FT DISULFID 520 533
FT DISULFID 591 605
FT DISULFID 643 648
FT METAL 78 78
FT METAL 110 110
FT METAL 210 210
FT METAL 271 271
FT METAL 413 413
FT METAL 451 451
FT METAL 544 544
FT METAL 613 613
FT BINDING 139 139
FT BINDING 481 481
FT CARBOHYD 118 118
FT CARBOHYD 494 494
FT CONFLICT 1 2
FT CONFLICT 25 25
FT CONFLICT 82 82
FT CONFLICT 359 359
FT CONFLICT 382 382
FT CONFLICT 449 449
FT CONFLICT 629 629
SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;

Query Match 64.3%; Score 45; DB 1; Length 707;
Best Local Similarity 63.6%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRNMKV 11
DB 37 CLRWQNMKV 47

RESULT 5
ID DYHB CHLRE STANDARD; PRT; 4568 AA.
AC Q39565;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dynein beta chain, flagellar outer arm.
GN ODA4 OR ODA-4 OR SUP1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=219E;
RX MEDLINE=94274778; PubMed=8006077;
RA Mitchell D.R., Brown K.S.;
RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy
chain genes."
RL J. Cell Sci. 107:635-644 (1994).
CC -!- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND
FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.
CC DYNEIN HAS ATPASE ACTIVITY.
CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND
GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.

```

```

CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J02963; AAA19956.1; -.
DR InterPro; IPR004273; Dynein heavy.
DR Pfam; PF03028; Dynein heavy; 1.
KW Motor protein; Microtubules; Dynein; ATP-binding; Flagella;
KW Coiled coil.
FT DOMAIN 277 293 COILED COIL (POTENTIAL).
FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).
FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).
FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).
FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).
FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).
FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).
FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).
FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).
FT NP_BIND 1919 1926 ATP (POTENTIAL).
FT NP_BIND 2202 2209 ATP (POTENTIAL).
FT NP_BIND 2530 2537 ATP (POTENTIAL).
FT NP_BIND 2879 2886 ATP (POTENTIAL).
SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;

Query Match 61.4%; Score 43; DB 1; Length 4568;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRNMKV 11
DB 1852 CFQWQSLRYI 1862

RESULT 6
ID TRFL HORSE STANDARD; PRT; 695 AA.
AC O77811;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin) (Fragment).
GN LTF.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "cDNA sequence of mare lactoferrin."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RC TISSUE=Milk;
RX MEDLINE=99296631; PubMed=10366507;
RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A
resolution."
RL J. Mol. Biol. 289:303-317 (1999).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AJ010930; CAA09407.1; --  
 CC PDB; 1BUX; 02-DEC-98.  
 CC PDB; 1B7U; 02-FEB-99.  
 CC PDB; 1B7Z; 02-FEB-99.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN 1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN 2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN 3; 1.  
 CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal; 3D-structure.

FT NON TER 1 1  
 FT SIGNAL <1 6  
 FT CHAIN 7 695 LACTOTRANSFERRIN.  
 FT REPEAT 7 350 1.  
 FT REPEAT 351 695 2.  
 FT DISULFID 15 51  
 FT DISULFID 25 42  
 FT DISULFID 121 204  
 FT DISULFID 163 179  
 FT DISULFID 166 189  
 FT DISULFID 176 187  
 FT DISULFID 237 251  
 FT DISULFID 354 386  
 FT DISULFID 364 377  
 FT DISULFID 411 690  
 FT DISULFID 431 653  
 FT DISULFID 463 538  
 FT DISULFID 487 681  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66 66 IRON 1 (BY SIMILARITY).  
 FT METAL 98 98 IRON 1 (BY SIMILARITY).  
 FT METAL 198 198 IRON 1 (BY SIMILARITY).  
 FT METAL 259 259 IRON 1 (BY SIMILARITY).  
 FT METAL 401 401 IRON 2 (BY SIMILARITY).  
 FT METAL 439 439 IRON 2 (BY SIMILARITY).  
 FT METAL 532 532 IRON 2 (BY SIMILARITY).  
 FT METAL 601 601 ANION (BY SIMILARITY).  
 FT BINDING 127 127 ANION (BY SIMILARITY).  
 FT BINDING 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 58.6%; Score 41; DB 1; Length 695;  
 Best Local Similarity 63.6%; Pred. No. 7;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQQRNMRKV 11  
 Db 25 CAKQRNMRKV 35

## RESULT 7

ID DCOR PANRE STANDARD; PRT; 435 AA.  
 AC P49725;  
 DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).  
 GN ODC.  
 OS Panagrellus redivivus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
 OX NCBI\_TaxID=6233;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95290001; PubMed=7772052;  
 RT von Besser H., Niemann G., Domdey B., Walter R.D.;  
 RT "Molecular cloning and characterization of ornithine decarboxylase  
 RT cDNA of the nematode Panagrellus redivivus.";  
 RL Biochem. J. 308:635-640(1995).  
 CC -|- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).  
 CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -|- PATHWAY: FIRST STEP AND ALSO THE RATE-LIMITING STEP IN THE PATHWAY  
 CC OF POLYAMINE BIOSYNTHESIS.  
 CC -|- SUBUNIT: HOMODIMER.  
 CC -|- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine  
 CC DECARBOXYLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X82199; CAA57683.1; --  
 CC EMBL; X95719; CAA65024.1; --  
 CC HSPP; P07805; 1P3T.  
 CC InterPro; IPR000183; Decarboxylase2.  
 CC Pfam; PF00278; Orn DAP Arg dec; 1.  
 CC Pfam; PF02784; Orn Arg dec N; 1.  
 CC PRINTS; PR01179; ODADCRBLASE.  
 CC PROSITE; PS00878; ODR\_DC\_2\_1; 1.  
 CC PROSITE; PS00879; ODR\_DC\_2\_2; 1.  
 CC Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis.  
 FT BINDING 76 76 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT ACT SITE 368 368 BY SIMILARITY.  
 SQ SEQUENCE 435 AA; 47110 MW; BBB093C1EF7FEFA4 CRC64;

Query Match 55.7%; Score 39; DB 1; Length 435;  
 Best Local Similarity 70.0%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QWQRNMRKVA 12  
 Db 61 QWQRNMRVA 70

## RESULT 8

RPOB LIBAF  
 ID RPOB LIBAF STANDARD; PRT; 146 AA.  
 AC P41187;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase  
 DE beta chain) (RNA polymerase beta subunit) (Fragment).  
 GN RPOB.  
 OS Liberibacter africanus (Liberibacter africanum).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Candidatus Liberibacter.  
 OX NCBI\_TaxID=34020;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Nelspruit;  
 RA Planet P., Jagoueix S., Bove J.M., Garnier M.;  
 RT "Detection and characterization of the African Citrus Greening

RT Liberoacter by amplification, cloning and sequencing of the rplKAJL-  
 RT ipoBC operon.";  
 RI Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 CC BETA' CHAIN.  
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U09675; AAA19557.1; -;  
 DR InterPro; IPR001572; RNA\_pol\_B.  
 DR Pfam; PF00562; RNA\_pol\_B; 1.  
 DR PROSITE; PS01166; RNA\_POL\_BETA; PARTIAL.  
 KW Transferase; Transcription; DNA-directed RNA polymerase.  
 FT NON\_TER 146 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FDB943 CRC64;  
 Query Match 54.3%; Score 38; DB 1; Length 146;  
 Best Local Similarity 60.0%; Pred. No. 5.4;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWQNRWK 10  
 DB 10 CVQWGRGARK 19  
 RESULT 9  
 VIA\_BMW STANDARD; PRT; 966 AA.  
 AC Q00020;  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1A protein [includes: Helicase, Methyltransferase].  
 OS Broad bean mottle virus.  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;  
 CC Bromovirus.  
 OX NCBI\_TaxID=12301;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Badwin;  
 RX MEDLINE=92074218; PubMed=1962437;  
 RA Dzanott A.M., Bujarski J.J.;  
 RT "The nucleotide sequence and genome organization of the RNA-1 segment  
 RT in two bromoviruses: broad bean mottle virus and cowpea chlorotic  
 RT mottle virus".  
 RL Virology 185:553-562 (1991).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.  
 CC CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE  
 CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA  
 CC CAPPING.  
 CC -!- SIMILARITY: TO 1A PROTEIN FROM CCNV, CMV, PSV AND TAV.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; MG5138; AAA42740.1; -;  
 DR PIR; A41699; PLVBVB.  
 DR InterPro; IPR002588; V\_methyltransf.  
 DR InterPro; IPR000606; Viral\_helicase1.  
 DR Pfam; PF01443; Viral\_helicase1; 1.  
 DR Pfam; PF01660; Vmethyltransf; 1.  
 KW Helicase; ATP-binding; Transferase; Methyltransferase.  
 FT NP\_BIND 690 697  
 SQ SEQUENCE 966 AA; 109621 MW; DF592681D7231C8D CRC64;  
 Query Match 54.3%; Score 38; DB 1; Length 966;  
 Best Local Similarity 50.0%; Pred. No. 37;  
 Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;  
 QY 1 CFQ-----WQNRWKVA 12  
 DB 347 CFKKNKDWTEENRWSVA 362  
 RESULT 10  
 DHOM\_METCL STANDARD; PRT; 435 AA.  
 ID DHOM\_METCL  
 AC P37143;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Homoserine dehydrogenase (EC 1.1.1.3) (Hdh).  
 GN HOM.  
 OS Methylobacillus glycogenes.  
 CC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;  
 CC Methylobacillus.  
 OX NCBI\_TaxID=406;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 21276;  
 RX MEDLINE=94161493; PubMed=8117070;  
 RA Motoyama H., Maki K., Anazawa H., Ishino S., Teshiba S.;  
 RT "Cloning and nucleotide sequences of the homoserine dehydrogenase  
 RT genes (hom) and the threonine synthase genes (thrc) of the Gram-  
 RT negative obligate methylotroph Methylobacillus glycogenes.";  
 RL Appl. Environ. Microbiol. 60:111-119 (1994).  
 CC -!- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-  
 CC semialdehyde + NAD(P)H.  
 CC -!- PATHWAY: THIRD STEP IN CONVERSION OF L-ASPARTATE TO HOMOSERINE.  
 CC HOMOSERINE PARTICIPATES IN THE BIOSYNTHESIS OF THREONINE AND THEN  
 CC ISOLEUCINE AND IN THE BIOSYNTHESIS OF METHIONINE.  
 CC -!- SIMILARITY: BELONGS TO THE HOMOSERINE DEHYDROGENASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D14070; BAA40414.1; -;  
 DR InterPro; IPR001342; Homoserine\_dh.  
 DR InterPro; IPR005106; NAD\_binding\_3.  
 DR Pfam; PF00742; Homoserine\_dh; 1.  
 DR Pfam; PF03447; NAD\_binding\_3; 1.  
 DR PROSITE; PS01042; HOMOSER\_DHGENASE; FALSE NEG.  
 KW Oxidoreductase; NADP; Threonine biosynthesis; Isoleucine biosynthesis;  
 KW Methionine biosynthesis.  
 FT NP\_BIND 9 16  
 SQ SEQUENCE 435 AA; 48226 MW; 58468B0E7A81ACAB CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 435;  
 Best Local Similarity 66.7%; Pred. No. 25;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WQNRWKVA 12  
 |||: |||

```

Db 111 WORDFRVA 119

RESULT 11
ID YKYL CAEEL STANDARD; PRT; 455 AA.
AC Q19910;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F30B5.4 in chromosome IV.
GN F30B5.4.
CS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Miller N., Bradshaw H.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE OKL38 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90903; BAAJ17147.1; -
CC InterPro; IPR004147; ABC1.
CC InterPro; IPR000719; Euk_pkinase.
CC Pfam; PF03109; ABC1; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 566 AA; 65078 MW; 0C50CC04509FDCB3 CRC64;

Query Match 52.9%; Score 37; DB 1; Length 566;
Best Local Similarity 61.5%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 2 FQWQR--NMKVA 12
Db 469 FQWQRLENMLSLA 481
||||| | | |
||| | | |

RESULT 13
LOLB VIBCH STANDARD; PRT; 211 AA.
ID LOLB VIBCH STANDARD; PRT; 211 AA.
AC P57070;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer-membrane lipoprotein lolB precursor.
GN LOLB OR VC2181.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi L., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: PLAYS A CRITICAL ROLE IN THE INCORPORATION OF
CC LIPOPROTEINS IN THE OUTER MEMBRANE AFTER THEY ARE RELEASED BY
CC THE LOLA PROTEIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (By similarity).
CC
CC -1- SIMILARITY: BELONGS TO THE LOLB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004289; AAF95326.1; -
CC TIGR; VC2181; -
CC InterPro; IPR004565; lolB.
CC Pfam; PF03550; lolB; 1.
CC Chaperone; Outer membrane; Lipoprotein; Transport; Protein transport;
KW

```

KW Signal; Complete proteome.  
 FT SIGNAL 1 26 BY SIMILARITY.  
 FT CHAIN 27 211 OUTER-MEMBRANE LIPOPROTEIN LOLB.  
 FT LIPID 27 27 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 SQ SEQUENCE 211 AA; 24379 MW; FIEF70858484177E CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 211;  
 Best Local Similarity 45.5%; Pred. No. 18;  
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 FQWQNRKVA 12  
 Db 71 FQWQSPQKLS 81  
 RESULT 14  
 VIP\_SIVS4  
 ID\_VIF\_SIVS4 STANDARD; PRT; 214 AA.  
 AC P12505;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-NOV-1990 (Rel. 16, Last annotation update)  
 DE Virion infectivity factor (SOR protein) (Q protein).  
 GN VIF.  
 OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11737;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89262053; PubMed=2786147;  
 RA Hirsch V.M., Oimstead R.A., Murphy-Corb M., Purcell R.H.,  
 RA Johnson P.R.;  
 RT Nature 339:389-392(1989)  
 RL "An African primate lentivirus (SIVsm) closely related to HIV-2.";  
 CC -!- FUNCTION: DETERMINES VIRUS INFECTIVITY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X14307; CAA32484.1; -  
 CC PIR; S07989; S07989  
 CC HIV; X14307; VIF\$SMH4.  
 CC InterPro; IPR000475; Viral\_infect.  
 CC Pfam; PF00559; Vif; 1.  
 CC PRINTS; PR00349; VIRIONINFECT.  
 CC ProDom; PD000063; Viral\_infect; 1.  
 KW AIDS  
 SQ SEQUENCE 214 AA; 25140 MW; 9BC5884EC454BF3D CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 214;  
 Best Local Similarity 75.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QWQNRK 10  
 Db 170 QWRNRNK 177  
 RESULT 15  
 FL3L\_HUMAN  
 ID\_FL3L\_HUMAN STANDARD; PRT; 235 AA.  
 AC P49771;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3  
 DE ligand) (Flt3L).  
 GN FLT3LG.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94195428; PubMed=8145851;  
 RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,  
 RA Bazan J.F., Kastelein R., Hudak S., Wagner J., Mattson J., Luh J.,  
 RA Duda G., Martina N., Peterson D., Menon S., Shanafelt A.,  
 RA Muench M., Kelnner G., Nankawa R., Rennick D., Roncarolo M.G.,  
 RA Zlotnik A., Rosnet O., Dubreuil P., Birnbaum D., Lee F.;  
 RT "Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of  
 RT haematopoietic stem cells and is encoded by variant RNAs.";  
 RT Nature 368:643-648(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94235842; PubMed=8180375;  
 RA Lyman S.D., James L., Johnson L., Brasel K., de Vries P.,  
 RA Escobar S.S., Downey H., Splett R.R., Beckmann M.P., McKenna H.J.;  
 RT "Cloning of the human homologue of the murine flt3 ligand: a growth  
 RT factor for early hematopoietic progenitor cells.";  
 RL Blood 83:2795-2801(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=96032581; PubMed=7566977;  
 RA Lyman S.D., Stocking K., Davison B., Fletcher F., Johnson L.,  
 RA Escobar S.;  
 RT "Structural analysis of human and murine flt3 ligand genomic loci.";  
 RT Oncogene 11:1165-1172(1995).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=20343011; PubMed=10881197;  
 RA Savvides S.N., Boone T., Karplus P.A.;  
 RT "Flt3 ligand structure and unexpected commonalities of helical  
 RT bundles and cysteine knots.";  
 RL Nat. Struct. Biol. 7:486-491(2000).  
 CC -!- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC  
 CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING  
 CC FACTORS AND INTERLEUKINS.  
 CC -!- SUBUNIT: Homodimer (isoform 2).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);  
 CC secreted (isoform 2).  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a 1/membrane-bound (shown here)  
 CC and 2/soluble; are produced by alternative splicing.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U04806; AAA17999.1; -  
 CC EMBL; U03858; AAA19825.1; -  
 CC EMBL; U29874; AAA90949.1; -  
 CC EMBL; U29874; AAA90950.1; -  
 CC PDB; 1ETE; 09-JUN-00.  
 CC Genew; HGNC:3766; FLT3LG.  
 CC MIM; 600007; -  
 CC InterPro; IPR004213; Flt3 lig.  
 CC Pfam; PF02947; flt3 lig; 1.  
 KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal;  
 3D-structure.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 235 SL CYTOKINE.  
 FT DOMAIN 27 184 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 185 205 POTENTIAL.  
 FT DOMAIN 206 235 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 30 111  
 FT DISULFID 70 153  
 FT DISULFID 119 158

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 161 178 DSSTLPFPWSPRPLEATA -> VETVFRKVSQGLDLLTS  
 (IN ISOFORM 2).  
 FT VARSPLIC 179 235 MISSING (IN ISOFORM 2).  
 FT CONFLICT 72 72 G -> A (IN REF. 1).  
 SQ SEQUENCE 235 AA; 26416 MW; 73B95BF693B4CECF CRC64;

Query Match 51.4%; Score 36; DB 1; Length 235;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFWQORNRK 10  
 Db 204 CLHWQTRRR 213

Search completed: February 21, 2003, 07:27:54  
 Job time : 5.6 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107b-79  
Perfect score: 70  
Sequence: 1 CFQWQNRKVA 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	94.3	711	4	Q8TCD2
2	57	81.4	38	4	Q9UCY5
3	51	72.9	33	6	Q9TR80
4	42	60.0	298	16	Q8YP77
5	41	58.6	109	15	Q9YQC1
6	41	58.6	109	15	Q9YQC0
7	41	58.6	109	15	Q9YQB9
8	41	58.6	109	15	Q9YQB8
9	41	58.6	109	15	Q9YJ17
10	41	58.6	109	15	Q9YJ32
11	41	58.6	148	10	Q9XHP1
12	41	58.6	500	10	Q41383
13	40	57.1	81	15	Q90863
14	40	57.1	329	12	Q9QB73
15	40	57.1	377	12	Q91MQ5
16	40	57.1	381	12	Q9DHK5

17	40	57.1	469	9	Q38115
18	39	55.7	105	10	Q9XFDS
19	39	55.7	121	16	Q8U6K3
20	39	55.7	279	16	Q8XSE2
21	39	55.7	306	4	Q8TAX2
22	39	55.7	372	10	Q81653
23	39	55.7	459	4	Q9NZW0
24	39	55.7	460	4	Q9NZW3
25	39	55.7	466	4	Q9NUS2
26	39	55.7	488	10	Q8S934
27	39	55.7	632	4	Q94937
28	39	55.7	932	5	Q19153
29	38	54.3	81	15	Q90884
30	38	54.3	205	8	Q98RR2
31	38	54.3	208	15	Q8US46
32	38	54.3	282	16	Q98Q19
33	38	54.3	294	11	Q9DCU6
34	38	54.3	341	11	Q8R2A4
35	38	54.3	376	12	Q9Q8X2
36	38	54.3	376	12	Q9Q8J4
37	38	54.3	397	16	Q9XAK9
38	38	54.3	464	16	Q8YP95
39	38	54.3	489	16	P96223
40	38	54.3	515	10	Q22185
41	38	54.3	543	10	Q22188
42	38	54.3	570	10	Q8S487
43	38	54.3	2186	5	Q9N906
44	37	52.9	87	15	Q90AB9
45	37	52.9	109	15	Q9YJF7

#### ALIGNMENTS

RESULT 1  
Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the ENBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -.  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 94.3%; Score 66; DB 4; Length 711;  
Best Local Similarity 100.0%; Pred. No. 0.0006;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRKVK 11  
Db 39 CFQWQNRKVK 49

RESULT 2  
Q9UCY5 PRELIMINARY; PRT; 38 AA.  
ID Q9UCY5  
AC Q9UCY5;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 RT seminal plasma";  
 RL Jpn. J. Legal Med. 49:281-293(1995).  
 DR HSP; P02788; 1BKA.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490BEBDBDEB CRC64;  
 Query Match 81.4%; Score 57; DB 4; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FQWQNRKRV 11  
 Db 21 FQWQNRKRV 30  
 RESULT 3  
 Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa;  
 OC Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=9512729; PubMed=7827104;  
 RA Qian Z.Y., Jolles P., Migliore-Samou D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32(1995).  
 DR HSP; O77698; 1CE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;  
 Query Match 72.9%; Score 51; DB 6; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.014;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQNRKRV 11  
 Db 19 CFQWQNRKRV 29  
 RESULT 4  
 Q8YP77 PRELIMINARY; PRT; 298 AA.  
 AC Q8YP77;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein ALR4323.  
 GN ALR4323.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP03596; BAB76022.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 298 AA; 34513 MW; 64036B852299A9F CRC64;  
 Query Match 60.0%; Score 42; DB 16; Length 298;  
 Best Local Similarity 77.8%; Pred. No. 7.2;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FQWQNRKRV 10  
 Db 163 FQWQNRKRV 171  
 RESULT 5  
 Q9YQCI PRELIMINARY; PRT; 109 AA.  
 AC Q9YQCI;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Envelope protein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98440505; PubMed=9765386;  
 RA Pasquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,  
 RA Puel J., Izopet J.;  
 RT "Molecular evidence for mother-to-child transmission of multiple  
 RT variants by analysis of RNA and DNA sequences of human  
 RT immunodeficiency virus";  
 RL J. Virol. 72:8493-8501(1998).  
 DR EMBL; AJ008835; CAA08297.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 109  
 FT NON\_TER 109  
 SQ SEQUENCE 109 AA; 12015 MW; EE0CA3E4A0A0D0EB CRC64;  
 Query Match 58.6%; Score 41; DB 15; Length 109;  
 Best Local Similarity 70.0%; Pred. No. 3.8;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 QWQNRKVA 12  
 Db 78 QWQNRKVA 87  
 RESULT 6  
 Q9YQCI PRELIMINARY; PRT; 109 AA.  
 AC Q9YQCI;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Envelope protein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98440505; PubMed=9765386;  
 RA Pasquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,  
 RA Puel J., Izopet J.;

RT "Molecular evidence for mother-to-child transmission of multiple  
RT variants by analysis of RNA and DNA sequences of human  
RT immunodeficiency virus.";

RL J. Virol. 72:8493-8501(1998).  
DR EMBL; AJ008839; CAA08301.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 12012 MW; FB18E5E4B546FB0 CRC64;

Query Match 58.6%; Score 41; DB 15; Length 109;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQNNMKVA 12  
|||:||||  
Db 78 QWNRTLKVA 87

## RESULT 7

Q9YQB8  
ID Q9YQB8 PRELIMINARY; PRT; 109 AA.  
AC Q9YQB8;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE Envelope protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98440505; PubMed=9765386;  
RA Pasquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,  
RA Puel J., Izopet J.;  
RT "Molecular evidence for mother-to-child transmission of multiple  
RT variants by analysis of RNA and DNA sequences of human  
RT immunodeficiency virus.";

RL J. Virol. 72:8493-8501(1998).  
DR EMBL; AJ008843; CAA08305.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 12053 MW; 8217D3EA2DFD4C4B CRC64;  
Query Match 58.6%; Score 41; DB 15; Length 109;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQNNMKVA 12  
|||:||||  
Db 78 QWNRTLKVA 87

## RESULT 8

Q9YQB8  
ID Q9YQB8 PRELIMINARY; PRT; 109 AA.  
AC Q9YQB8;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE Envelope protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=98440505; PubMed=9765386;  
RA Pasquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,  
RA Puel J., Izopet J.;  
RT "Molecular evidence for mother-to-child transmission of multiple  
RT variants by analysis of RNA and DNA sequences of human  
RT immunodeficiency virus.";

RL J. Virol. 72:8493-8501(1998).  
DR EMBL; AJ008844; CAA08306.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 12066 MW; 81E41CCEBD645E4B CRC64;

Query Match 58.6%; Score 41; DB 15; Length 109;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQNNMKVA 12  
|||:||||  
Db 78 QWNRTLKVA 87

## RESULT 9

Q9YJ17  
ID Q9YJ17 PRELIMINARY; PRT; 109 AA.  
AC Q9YJ17;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-OCT-2001 (TREMELrel. 18, Last annotation update)  
DE Envelope protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98440505; PubMed=9765386;  
RA Pasquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,  
RA Puel J., Izopet J.;  
RT "Molecular evidence for mother-to-child transmission of multiple  
RT variants by analysis of RNA and DNA sequences of human  
RT immunodeficiency virus type 1.";

RL J. Virol. 72:8493-8501(1998).  
DR EMBL; AJ008843; CAA08303.1; -.  
DR EMBL; AJ008836; CAA08298.1; -.  
DR EMBL; AJ008837; CAA08299.1; -.  
DR EMBL; AJ008838; CAA08300.1; -.  
DR EMBL; AJ008840; CAA08302.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 12042 MW; EE0CA3E4B04EDE1B CRC64;  
Query Match 58.6%; Score 41; DB 15; Length 109;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQNNMKVA 12  
|||:||||  
Db 78 QWNRTLKVA 87

## RESULT 10

Q9YJ2  
ID Q9YJ2 PRELIMINARY; PRT; 109 AA.  
AC Q9YJ2;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-OCT-2001 (TREMELrel. 18, Last annotation update)

```

DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
SC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98440505; PubMed=9765386;
RA Paquier C.J.M., Cayrou C., Blancher A., Berrebi A., Tricoire J.,
RA Fuet J., Izopet J.;
RT "Molecular evidence for mother-to-child transmission of multiple
RT variants by analysis of RNA and DNA sequences of human
RT immunodeficiency virus type 1.";
RL J. Virol. 72:8493-8501(1998).
DR EMBL; AJ009061; CAA08523.1; -.
DR EMBL; AJ008842; CAA08304.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein.
FT NON TER 1
FT NON TER 109
SQ SEQUENCE 109 AA; 12028 MW; E80CA3F708754E1B CRC64;

Query Match 58.6%; Score 41; DB 15; Length 109;
Best Local Similarity 70.0%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QWNRKVA 12
Db 78 QWNRKVA 87
|||:|||||
|||:|||||

RESULT 11
QXHP1 PRELIMINARY; PRT; 148 AA.
ID Q9XHP1
AC Q9XHP1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2S albumin.
OS Sesamum indicum (Oriental sesame) (Gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20074970; PubMed=10606554;
RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;
RT "Molecular cloning of 15S globulin and 2S albumin, the two major seed
RT storage proteins in sesame.";
RL J. Agric. Food Chem. 47:4932-4938(1999).
DR EMBL; AF091841; AAD42943.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000617; Napin.
DR InterPro; IPR001768; Try/amy1_inhbr.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00496; NAPIN.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 58.6%; Score 41; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQNRMR 9
Db 54 CFQWQNRMR 62
|||:|||||
|||:|||||

RESULT 12
Q41383 PRELIMINARY; PRT; 500 AA.
ID Q41383;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LEAF;
RA Baur B., Winter K., Fischer K., Dietz K.;
RT "Molecular cloning and characterization of several protein kinases
RT from plants.";
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z30330; CAA82991.1; -.
DR HSP; P05132; ICTP.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 500 AA; 57996 MW; 5579A056AC357C07 CRC64;

Query Match 58.6%; Score 41; DB 10; Length 500;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WQNRKVA 12
Db 268 WQNRKVA 276
|||:|||||
|||:|||||

RESULT 13
Q90863 PRELIMINARY; PRT; 81 AA.
ID Q90863
AC Q90863;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein Gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
SC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NJS182;
RA Brandful J.A.M., Ampofo W.K., Janssens W., Adu-Sarkodie Y.,
RA Apeagyei F., Anyomi F., Aidoo S., Barnor J.S., Yamamoto N.,
RA Ishikawa K., Sata T., Kurata T.;
RT "Genetic and phylogenetic analysis of HIV-1 strains from Southern
RT Ghana.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ225659; CAA12541.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON TER 1
FT NON TER 81
SQ SEQUENCE 81 AA; 9138 MW; 2D43DCD554295572 CRC64;

```

Query Match 57.1%; Score 40; DB 15; Length 81;  
Best Local Similarity 60.0%; Pred. No. 4.2;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QMORNKVA 12  
: : : : :  
Db 69 EWKENLRKVA 78

## RESULT 14

QSQB73  
ID Q9QB73 PRELIMINARY; PRT; 329 AA.  
AC Q9QB73;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Yb-DLL protein (Fragment).  
GN Yb-DLL.  
OS Yaba monkey tumor virus.  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Yatapoxvirus.  
CX NCBI\_TaxID=38804;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Amano a., Miyamura T.;  
RT "DNA sequence of Yaba virus BamHI-D fragment."  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB018404; BAA88879.1; -  
DR InterPro; IPR004251; DUF230.  
DR Pfam; PF03003; DUF230; 1.  
FT NON TER 329 329  
SQ SEQUENCE 329 AA; 37969 MW; 03511EB8F7582FC2 CRC64;

Query Match 57.1%; Score 40; DB 12; Length 329;  
Best Local Similarity 50.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWORNKVA 12  
: : : : :  
Db 186 CLKWLRTKEKIA 196

## RESULT 15

Q91MQ5  
ID Q91MQ5 PRELIMINARY; PRT; 377 AA.  
AC Q91MQ5;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE LSPV108 putative myristylated membrane protein.  
GN LSPV108.  
OS lumpy skin disease virus.  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Capripoxvirus.  
CX NCBI\_TaxID=59509;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEETHLING 2490;  
RX MEDLINE=21329495; PubMed=11435593;  
RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
RT "Genome of lumpy skin disease virus."  
RL J. Virol. 75:7122-7130(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEETHLING 2490;  
RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF325528; AAK85069.1; -  
DR InterPro; IPR004251; DUF230.  
DR Pfam; PF03003; DUF230; 1.  
SQ SEQUENCE 377 AA; 43937 MW; FA3B4316D8B28D3B CRC64;

Query Match 57.1%; Score 40; DB 12; Length 377;

Best Local Similarity 50.0%; Pred. No. 22;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWORNKVA 12  
: : : : :  
Db 185 CLKWLRTKEKIA 196

Search completed: February 21, 2003, 07:44:33  
Job time : 20.8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 Seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107B-80  
Perfect score: 64  
Sequence: 1 CFQLQRNKRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

- 1: /SID32/gcgdata/geneseq/geneseq-embl/AA1980.DAT:\*
- 2: /SID32/gcgdata/geneseq/geneseq-embl/AA1981.DAT:\*
- 3: /SID32/gcgdata/geneseq/geneseq-embl/AA1982.DAT:\*
- 4: /SID32/gcgdata/geneseq/geneseq-embl/AA1983.DAT:\*
- 5: /SID32/gcgdata/geneseq/geneseq-embl/AA1984.DAT:\*
- 6: /SID32/gcgdata/geneseq/geneseq-embl/AA1985.DAT:\*
- 7: /SID32/gcgdata/geneseq/geneseq-embl/AA1986.DAT:\*
- 8: /SID32/gcgdata/geneseq/geneseq-embl/AA1987.DAT:\*
- 9: /SID32/gcgdata/geneseq/geneseq-embl/AA1988.DAT:\*
- 10: /SID32/gcgdata/geneseq/geneseq-embl/AA1989.DAT:\*
- 11: /SID32/gcgdata/geneseq/geneseq-embl/AA1990.DAT:\*
- 12: /SID32/gcgdata/geneseq/geneseq-embl/AA1991.DAT:\*
- 13: /SID32/gcgdata/geneseq/geneseq-embl/AA1992.DAT:\*
- 14: /SID32/gcgdata/geneseq/geneseq-embl/AA1993.DAT:\*
- 15: /SID32/gcgdata/geneseq/geneseq-embl/AA1994.DAT:\*
- 16: /SID32/gcgdata/geneseq/geneseq-embl/AA1995.DAT:\*
- 17: /SID32/gcgdata/geneseq/geneseq-embl/AA1996.DAT:\*
- 18: /SID32/gcgdata/geneseq/geneseq-embl/AA1997.DAT:\*
- 19: /SID32/gcgdata/geneseq/geneseq-embl/AA1998.DAT:\*
- 20: /SID32/gcgdata/geneseq/geneseq-embl/AA1999.DAT:\*
- 21: /SID32/gcgdata/geneseq/geneseq-embl/AA2000.DAT:\*
- 22: /SID32/gcgdata/geneseq/geneseq-embl/AA2001.DAT:\*
- 23: /SID32/gcgdata/geneseq/geneseq-embl/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	12	21	AAV78080 Human lactoferrin
2	59	92.2	12	21	AAV78071 Human lactoferrin
3	58	90.6	12	21	AAV78038 Human lactoferrin
4	58	90.6	12	21	AAV78046 Human lactoferrin
5	58	90.6	12	21	AAV78047 Human lactoferrin
6	58	90.6	12	21	AAV78096 Human lactoferrin
7	58	90.6	13	21	AAV78037 Human lactoferrin
8	58	90.6	13	21	AAV78048 Human lactoferrin
9	58	90.6	13	21	AAV78049 Human lactoferrin
10	58	90.6	14	21	AAV78036 Human lactoferrin

11	58	90.6	14	21	AAV78050 Human lactoferrin
12	58	90.6	14	21	AAV78051 Human lactoferrin
13	58	90.6	15	17	AAE98554 Peptide for anti-u
14	58	90.6	15	21	AAV78035 Human lactoferrin
15	58	90.6	15	21	AAV78062 Human lactoferrin
16	58	90.6	15	21	AAV78063 Human lactoferrin
17	58	90.6	16	21	AAV78031 Human lactoferrin
18	58	90.6	16	21	AAV78064 Human lactoferrin
19	58	90.6	16	21	AAV78065 Human lactoferrin
20	58	90.6	17	21	AAV78034 Human lactoferrin
21	58	90.6	17	21	AAV78066 Human lactoferrin
22	58	90.6	17	21	AAV78067 Human lactoferrin
23	58	90.6	18	15	AAE69352 Human lactoferrin
24	58	90.6	18	17	AAW13397 Advanced glycosyla
25	58	90.6	18	21	AAV78033 Human lactoferrin
26	58	90.6	19	21	AAV68867 Amino acid sequenc
27	58	90.6	19	21	AAV78032 Human lactoferrin
28	58	90.6	20	13	AAE21810 Anti microbial pep
29	58	90.6	20	14	AAE44841 Lactoferrin-relate
30	58	90.6	20	15	AAE48530 Lactoferrin derive
31	58	90.6	20	15	AAE48531 Lactoferrin derive
32	58	90.6	20	15	AAE57461 Lactoferrin derive
33	58	90.6	20	15	AAE57462 Lactoferrin derive
34	58	90.6	20	16	AAE84698 Bovine lactoferrin
35	58	90.6	20	16	AAE84699 Bovine lactoferrin
36	58	90.6	20	16	AAE80263 Anti-parasitic lac
37	58	90.6	20	16	AAE80264 Anti-parasitic lac
38	58	90.6	20	17	AAE98553 Peptide for anti-u
39	58	90.6	20	17	AAE91852 Lactoferrin-derive
40	58	90.6	20	17	AAW03045 Lactoferrin-derive
41	58	90.6	20	17	AAE90607 Lactoferrin-derive
42	58	90.6	20	17	AAE87621 Lactoferrin-derive
43	58	90.6	20	17	AAE87622 Lactoferrin-derive
44	58	90.6	20	18	AAW26150 Lactoferrin deriva
45	58	90.6	20	18	AAW14036 Anti-parasitic pep

## ALIGNMENTS

RESULT 1  
AAV78080  
ID AAV78080 standard; Peptide; 12 AA.  
XX AC AAV78080;  
XX DT 25-APR-2000 (first entry)  
XX DE Human lactoferrin derived peptide SEQ ID NO:80.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.

XX OS Homo sapiens.  
XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
SQ Query Match 100.0%; Score 64; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQLQRMNRKVR 12  
Db 1 CFQLQRMNRKVR 12

RESULT 2  
AAY78071  
ID AAY78071 standard; Peptide; 12 AA.  
XX  
AC AAY78071;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:71.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson IA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
DR WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 35; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
SQ Query Match 92.2%; Score 59; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00039;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLQRMNRKVR 12  
Db 1 CFQLQRMNRKVR 12

RESULT 3  
AAY78038  
ID AAY78038 standard; Peptide; 12 AA.  
XX  
AC AAY78038;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:38.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson IA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
DR WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 12; Page 70; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 90.6%; Score 58; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00059;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQQRNMRKVR 12  
 DB 1 CFQQRNMRKVR 12

RESULT 4  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 XX AAY78046;  
 AC AAY78046;  
 DT 25-APR-2000 (first entry)  
 DE Human lactoferrin derived peptide SEQ ID NO:46.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 90.6%; Score 58; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00059;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQQRNMRKVR 12  
 DB 1 CFQQRNMRKVR 12

RESULT 5  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.  
 XX AAY78047;  
 AC AAY78047;  
 DT 25-APR-2000 (first entry)  
 DE Human lactoferrin derived peptide SEQ ID NO:47.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 90.6%; Score 58; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00059;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 1 CFQLQNNMKVR 12  
 DB 1 CFQWQNNMKVR 12

## RESULT 6

AAV78096  
 ID AAV78096 standard; Peptide; 12 AA.

XX AAV78096;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:96.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

XX inflammations and tumors and for use in infant formula food

XX Claim 22; Page 38; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 90.6%; Score 58; DB 21; Length 12;

Best Local Similarity 83.3%; Pred. No. 0.00059;

Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQLQNNMKVR 12

DB 1 CFQLQNNMKVR 12

## RESULT 7

AAV78037

ID AAV78037 standard; Peptide; 13 AA.

XX AAV78037;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:37.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

XX inflammations and tumors and for use in infant formula food

XX Claim 12; Page 70; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 90.6%; Score 58; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.00064;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLQNNMKVR 12

DB 2 CFQWQNNMKVR 13

## RESULT 8

AAV78048

ID AAV78048 standard; Peptide; 13 AA.

XX AAV78048;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 74; 102pp; English.  
 XX  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 90.6%; Score 58; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.00064;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLOQNMKVR 12  
 DB ||| |||||  
 2 CFQWQNMKVR 13  
 RESULT 9  
 AAY78049  
 ID AAY78049 standard; Peptide; 13 AA.  
 XX  
 AC AAY78049;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:49.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.  
 XX  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 90.6%; Score 58; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.00064;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLOQNMKVR 12  
 DB ||| |||||  
 2 CFQWQNMKVR 13  
 RESULT 10  
 AAY78036  
 ID AAY78036 standard; Peptide; 14 AA.  
 XX  
 AC AAY78036;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:36.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.

```

XX PR 06-JUL-1998; 98SE-0002441.
XX PR 17-JUL-1998; 98SE-0002562.
XX PR 29-DEC-1998; 98SE-0004614.
XX PA (ASCI-) A+ SCI INVEST AB.
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX DR WPI; 2000-147388/13.
XX XX
XX PT New peptides used for treatment and prevention of infections,
XX PT inflammations and tumors and for use in infant formula food -
XX PS Claim 12; Page 69; 102pp; English.
XX XX
XX CC AAY78001 to AAY78100 represent peptides having sequences based on human
XX CC lactoferrin. The peptides are taken up in the intestine through
XX CC binding to specific lactoferrin receptors and are then transported
XX CC through the circulation. A medicinal product of the peptide or fragment
XX CC can be used for treating and/or prevention of infections (such as
XX CC urinary tract infections, colitis, and Candida infection on a mucosal
XX CC membrane), inflammations and/or tumors. The peptides can also be
XX CC used in food stuffs such as infant formula food. The peptides are also
XX CC fungicidal and bactericidal and may also be used as preservatives.
XX CC Even though native human lactoferrin have been shown to have desired
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they
XX CC cannot be used clinically on a broad basis because of high production
XX CC costs. Therefore, provision of peptides based on lactoferrin would
XX CC enable them to be used for the same purposes as lactoferrin at lower
XX CC cost.
XX XX
XX SQ Sequence 14 AA;
XX
Query Match 90.6%; Score 58; DB 21; Length 14;
Best Local Similarity 91.7%; Pred. No. 0.00069;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLQRNMRKVR 12
Db 3 CFQWRNMRKVR 14
||| |||||
RESULT 11
AAY78050
ID AAY78050 standard; Peptide; 14 AA.
AC AAY78050;
XX
XX DT 25-APR-2000 (first entry)
XX DE Human lactoferrin derived peptide SEQ ID NO:50.
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX KW urinary tract infection; colitis; Candida infection; fungicidal;
XX KW bactericidal; preservative.
XX OS Homo sapiens.
XX OS Synthetic.
XX XX
XX PN WO200001730-A1.
XX XX
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-SE01230.
XX XX
XX PR 06-JUL-1998; 98SE-0002441.
XX PR 17-JUL-1998; 98SE-0002562.
XX PR 29-DEC-1998; 98SE-0004614.
XX PA (ASCI-) A+ SCI INVEST AB.
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX DR WPI; 2000-147388/13.
XX XX
XX PT New peptides used for treatment and prevention of infections,
XX PT inflammations and tumors and for use in infant formula food -
XX XX

QY 1 CFQLQRNMRKVR 12
Db 3 CFQWRNMRKVR 14
||| |||||
RESULT 12
AAY78051
ID AAY78051 standard; Peptide; 14 AA.
AC AAY78051;
XX
XX DT 25-APR-2000 (first entry)
XX DE Human lactoferrin derived peptide SEQ ID NO:51.
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX KW urinary tract infection; colitis; Candida infection; fungicidal;
XX KW bactericidal; preservative.
XX OS Homo sapiens.
XX OS Synthetic.
XX XX
XX PN WO200001730-A1.
XX XX
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-SE01230.
XX XX
XX PR 06-JUL-1998; 98SE-0002441.
XX PR 17-JUL-1998; 98SE-0002562.
XX PR 29-DEC-1998; 98SE-0004614.
XX PA (ASCI-) A+ SCI INVEST AB.
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX DR WPI; 2000-147388/13.
XX XX
XX PT New peptides used for treatment and prevention of infections,
XX PT inflammations and tumors and for use in infant formula food -
XX XX

```

Claim 18; Page 75; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

SQ Sequence 14 AA;

Query Match 90.6%; Score 58; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 0.00069;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRMRKVR 12

Db 3 CFQWQNRMRKVR 14

RESULT 13

ID AAR98554  
 ID AAR98554 standard; Peptide; 15 AA.

XX AAR98554;

XX 12-NOV-1996 (first entry)

DE Peptide for anti-ulcer agent.

XX anti-ulcer agent; low toxicity; stable; heat-resistant.

OS Synthetic.

XX JP08143468-A.

XX 04-JUN-1996.

XX 17-NOV-1994; 94JP-0283869.

XX 17-NOV-1994; 94JP-0283869.

XX (MORG) MORINAGA MILK IND CO LTD.

XX WPI; 1996-318857/32.

XX Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble

PS Claim 1; Page 11; 11pp; Japanese.

XX AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.

SQ Sequence 15 AA;

Query Match 90.6%; Score 58; DB 17; Length 15;

Best Local Similarity 91.7%; Pred. No. 0.00074;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRMRKVR 12

Db 2 CFQWQNRMRKVR 13

RESULT 14

AAY78035

ID AAY78035 standard; Peptide; 15 AA.

XX AAY78035;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:35.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 12; Page 69; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 15 AA;

Query Match

Best Local Similarity 90.6%; Score 58; DB 21; Length 15;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRMRKVR 12

Db 4 CFQWQNRMRKVR 15

RESULT 15

AAY78062

ID AAY78062 standard; Peptide; 15 AA.

XX AAY78062;

AC

XX 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:62.  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1999; 98SE-0002441.  
PR 17-JUL-1999; 98SE-0002562.  
PR 29-DEC-1999; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI WPI; 2000-147388/13.  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 15; Page 81; 102pp; English.  
XX AA78001 to AA78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX SQ Sequence 15 AA;  
Query Match 90.6%; Score 58; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. NO. 0.00074;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CQQLQNNMKVR 12  
DB 4 CQQLQNNMKVR 15  
Search completed: February 21, 2003, 07:37:14  
Job time : 29.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 Seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-80

Perfect score: 64

Sequence: 1 CFQLQRNMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	90.6	18	1	US-08-204-487-3
2	58	90.6	18	2	US-08-485-948-8
3	58	90.6	18	2	US-08-628-380-8
4	58	90.6	18	2	US-08-475-055-8
5	58	90.6	20	1	US-07-755-161A-3
6	58	90.6	20	1	US-07-891-174-3
7	58	90.6	20	1	US-08-204-487-1
8	58	90.6	20	1	US-08-256-771-24
9	58	90.6	20	1	US-08-256-771-25
10	58	90.6	20	1	US-08-381-984-24
11	58	90.6	20	1	US-08-381-984-25
12	58	90.6	22	4	US-09-508-734-4
13	58	90.6	24	1	US-09-508-734-6
14	58	90.6	25	1	US-07-755-161A-10
15	58	90.6	25	1	US-07-891-174-10
16	58	90.6	25	1	US-08-204-487-7
17	58	90.6	29	4	US-09-508-734-8
18	58	90.6	36	1	US-07-755-161A-8
19	58	90.6	36	1	US-07-891-174-8
20	58	90.6	36	1	US-08-256-771-30
21	58	90.6	36	1	US-08-381-984-29
22	58	90.6	47	2	US-08-464-182A-6
23	58	90.6	47	2	US-08-406-271-6
24	58	90.6	50	2	US-08-693-274A-7
25	58	90.6	52	4	US-09-017-043A-3
26	58	90.6	53	2	US-08-464-182A-5
27	58	90.6	53	2	US-08-406-271-5

28	58	90.6	54	2	US-08-464-182A-2
29	58	90.6	54	2	US-08-406-271-2
30	58	90.6	694	3	US-08-724-586-2
31	58	90.6	694	4	US-09-421-632-2
32	58	90.6	694	4	US-09-932-190-2
33	58	90.6	705	2	US-08-655-640-2
34	58	90.6	708	2	US-08-655-640-4
35	58	90.6	711	1	US-08-154-019-4
36	58	90.6	711	1	US-08-461-333-4
37	58	90.6	711	3	US-08-464-167-4
38	58	90.6	711	3	US-09-158-313-4
39	58	90.6	711	3	US-08-476-798-4
40	55	85.9	711	1	US-08-145-681-2
41	55	85.9	711	1	US-08-250-308-2
42	55	85.9	711	1	US-08-453-703-2
43	55	85.9	711	2	US-08-456-106-2
44	55	85.9	711	3	US-08-456-108-2
45	55	85.9	711	4	US-09-265-577-2

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37) "

US-08-204-487-3

Query Match 90.6%; Score 58; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQIQRMNRKVR 12  
||| |||||  
Db 1 CFQIQRMNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/485,948  
; FILING DATE: JUNE 7, 1995  
; CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 90.6%; Score 58; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQIQRMNRKVR 12  
||| |||||  
Db 1 CFQIQRMNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 90.6%; Score 58; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQIQRMNRKVR 12  
||| |||||  
Db 1 CFQIQRMNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/495,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-CI, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;;  
US-08-475-055-8

Query Match 90.6%; Score 58; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQLQNNMKVR 12  
||| |||||  
Db 1 CFQWQNNMKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
;;  
US-07-755-161A-3

Query Match 90.6%; Score 58; DB 1; Length 20;



Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLOQNRKVR 12  
Db 2 CFQWQNRKVR 13

## RESULT 6

US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991

ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM:

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY: modified site

LOCATION: 2  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 19"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 19  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 2"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLOQNRKVR 12  
Db 2 CFQWQNRKVR 13

## RESULT 7

US-08-204-487-1  
; Sequence 1, Application US/08204487

; Patent No. 5565425

; GENERAL INFORMATION:

; APPLICANT: YAMAMOTO, NAOKI

; APPLICANT: NAKASHIMA, HIDEKI

; APPLICANT: MOSUCHI, WATARU

; APPLICANT: TANAKA, SHIGEKI

; APPLICANT: DOSAKO, SHUN'ICHI

; APPLICANT: KAWASAKI, YOSHIHIRO

; APPLICANT: UCHIDA, TOSHIKI

; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/204,487

; FILING DATE: 02-MAR-1994

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: CAMPBELL, PAULA A.

; REGISTRATION NUMBER: 32,503

; REFERENCE/DOCKET NUMBER: RJN-019

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQQRNMRKVR 12  
||| |||||  
Db 2 CFQQRNMRKVR 13

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"  
US-08-204-487-1

US-08-256-771-24

Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQQRNMRKVR 12  
||| |||||  
Db 2 CFQQRNMRKVR 13

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
US-08-256-771-25

Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQQRNMRKVR 12  
||| |||||  
Db 2 CFQQRNMRKVR 13

RESULT 10  
US-08-381-984-24

; Sequence 24, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
; US-08-381-984-24  
Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQLOQNMKVR 12  
DB 2 CFQWQNMKVR 13  
RESULT 11  
US-08-381-984-25  
; Sequence 25, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack

; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
; US-08-381-984-25  
Query Match 90.6%; Score 58; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQLOQNMKVR 12  
DB 2 CFQWQNMKVR 13  
RESULT 12  
US-09-508-734-4  
; Sequence 4, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; TITLE OF INVENTION: useful microorganism thereof  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KR99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 4

LENGTH: 22  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-4

Query Match 90.6%; Score 58; DB 4; Length 22;  
Best Local Similarity 91.7%; Pred. No. 0.00046;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNMNRKVR 12  
||| |||||  
DB 2 CFQWQNMNRKVR 13

## RESULT 13

US-09-508-734-6  
Sequence 6, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: KopatentIn 1.71  
SEQ ID NO 6  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-6

Query Match 90.6%; Score 58; DB 4; Length 24;  
Best Local Similarity 91.7%; Pred. No. 0.0005;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNMNRKVR 12  
||| |||||  
DB 3 CFQWQNMNRKVR 14

## RESULT 14

US-07-755-161A-10  
Sequence 10, Application US/07755161A  
Patent No. 5304633  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/755,161A  
FILING DATE: 19910905  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
US-07-755-161A-10  
Query Match 90.6%; Score 58; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 0.00052;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQLOQNMNRKVR 12  
||| |||||  
DB 4 CFQWQNMNRKVR 15

RESULT 15  
US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-10  
Query Match 90.6%; Score 58; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 0.00052;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CQLOQNNMKVR 12  
||| |||||  
Db 4 CQWQNNMKVR 15  
Search completed: February 21, 2003, 07:50:35  
Job time : 8.7 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-80  
Perfect score: 64  
Sequence: 1 CFQLOQNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	90.6	15	9	US-09-798-869-2
2	58	90.6	25	9	US-09-798-869-20
3	58	90.6	694	9	US-10-023-096-2
4	50	78.1	15	9	US-09-798-869-6
5	41	64.1	15	9	US-09-798-869-3
6	41	64.1	25	9	US-09-798-869-23
7	38	59.4	97	9	US-09-738-626-5597
8	36	56.2	360	9	US-09-738-626-5597
9	35	54.7	700	9	US-10-013-310-5
10	35	54.7	882	9	US-10-174-590-574
11	35	54.7	882	9	US-10-176-758-574
12	35	54.7	882	9	US-10-173-737-574
13	35	54.7	882	9	US-10-173-706-574
14	35	54.7	882	9	US-10-175-738-574
15	35	54.7	882	9	US-10-175-752-574
16	35	54.7	882	9	US-10-176-482-574
17	35	54.7	882	9	US-10-176-757-574
18	35	54.7	882	9	US-10-176-913-574
19	35	54.7	882	9	US-10-180-552-574

20 35 54.7 882 9 US-10-180-557-574 Sequence 574, App  
21 35 54.7 882 9 US-10-173-700-574 Sequence 574, App  
22 35 54.7 882 9 US-10-174-572-574 Sequence 574, App  
23 35 54.7 882 9 US-10-174-579-574 Sequence 574, App  
24 35 54.7 882 9 US-10-174-582-574 Sequence 574, App  
25 35 54.7 882 9 US-10-174-588-574 Sequence 574, App  
26 35 54.7 882 9 US-10-175-739-574 Sequence 574, App  
27 35 54.7 882 9 US-10-175-740-574 Sequence 574, App  
28 35 54.7 882 9 US-10-175-743-574 Sequence 574, App  
29 35 54.7 882 9 US-10-176-488-574 Sequence 574, App  
30 35 54.7 882 9 US-10-176-492-574 Sequence 574, App  
31 35 54.7 882 9 US-10-176-747-574 Sequence 574, App  
32 35 54.7 882 9 US-10-176-750-574 Sequence 574, App  
33 35 54.7 882 9 US-10-176-985-574 Sequence 574, App  
34 35 54.7 882 9 US-10-176-987-574 Sequence 574, App  
35 35 54.7 882 9 US-10-176-991-574 Sequence 574, App  
36 35 54.7 882 9 US-10-176-992-574 Sequence 574, App  
37 35 54.7 882 9 US-10-176-993-574 Sequence 574, App  
38 35 54.7 882 9 US-10-184-658-574 Sequence 574, App  
39 35 54.7 882 9 US-10-173-695-574 Sequence 574, App  
40 35 54.7 882 9 US-10-173-697-574 Sequence 574, App  
41 35 54.7 882 9 US-10-173-705-574 Sequence 574, App  
42 35 54.7 882 9 US-10-174-576-574 Sequence 574, App  
43 35 54.7 882 9 US-10-174-585-574 Sequence 574, App  
44 35 54.7 882 9 US-10-174-586-574 Sequence 574, App  
45 35 54.7 882 9 US-10-175-747-574 Sequence 574, App

## ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 90.6%; Score 58; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 6.7e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRKVR 12  
||| |||||  
DB 3 CFQLOQNRKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 90.6%; Score 58; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRKVR 12  
||| |||||  
DB 3 CFQWQNRKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 90.6%; Score 58; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0035;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNRKVR 12  
||| |||||  
DB 22 CFQWQNRKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINEJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 78.1%; Score 50; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLOQNRKVR 12  
||| |||||  
DB 3 CFQWQNRKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINEJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 64.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.1;

Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 0;  
QY 1 CFQLOQRMKRV 11  
|:|:|:|:|:|:  
DB 3 CYQWQRRMRKL 13

RESULT 6  
US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (VSTEIN REKDAL  
; APPLICANT: BALOUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 64.1%; Score 41; DB 9; Length 25;  
Best Local Similarity 63.8%; Pred. No. 0.17;  
Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 1 CFQLOQRMKRV 11  
|:|:|:|:|:|:  
DB 3 CYQWQRRMRKL 13

RESULT 7  
US-09-738-626-5597  
; Sequence 5597, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOL, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5597  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5597

Query Match 59.4%; Score 38; DB 9; Length 97;  
Best Local Similarity 70.0%; Pred. No. 2.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLOQRMKRV 10  
|:|:|:|:|:|:  
DB 73 CPELVQRMKRV 82

RESULT 8  
US-09-829-378-7  
; Sequence 7, Application US/09829378  
; Patent No. US20020170082A1  
; GENERAL INFORMATION:  
; APPLICANT: FOX, Timothy  
; APPLICANT: ALBERTSEN, Marc C.  
; TITLE OF INVENTION: GENE AFFECTING MALE FERTILITY IN PLANTS  
; FILE REFERENCE: 033229/0631  
; CURRENT APPLICATION NUMBER: US/09/829,378  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/340,684  
; PRIOR FILING DATE: 2001-06-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-829-378-7

Query Match 56.2%; Score 36; DB 9; Length 360;  
Best Local Similarity 54.5%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLOQRMKRV 11  
|:|:|:|:|:|:  
DB 331 CPELVQRMKRV 341

RESULT 9  
US-10-013-310-5  
; Sequence 5, Application US/10013310  
; Publication No. US20020192216A1  
; GENERAL INFORMATION:  
; APPLICANT: Lamb, Jonathon Robert  
; APPLICANT: Hoyne, Gerard Francis  
; APPLICANT: Dallman, Margaret Jane  
; TITLE OF INVENTION: Therapeutic Use  
; FILE REFERENCE: 674525-2003  
; CURRENT APPLICATION NUMBER: US/10/013,310  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/GB00/02191  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: UK 9913350.6  
; PRIOR FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: UK 9921953.7  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 700  
; TYPE: PRT  
; ORGANISM: House Mouse  
US-10-013-310-5

Query Match 54.7%; Score 35; DB 9; Length 700;  
Best Local Similarity 54.5%; Pred. No. 70;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLOQRMKRV 11  
|:|:|:|:|:|:  
DB 666 CEQVDRNRV 676



QY 1 CFOLQNRMKV 11  
Db 848 CEQVDRNIRRV 858

RESULT 12  
US-10-175-737-574  
; Sequence 574, Application US/10175737  
; Publication No. US20030013153A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C50  
; CURRENT APPLICATION NUMBER: US/10/175,737  
; CURRENT FILING DATE: 2002-06-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-737-574

Query Match 54.7%; Score 35; DB 9; Length 882;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLQNRMKV 11  
Db 848 CEQVDRNIRRV 858

RESULT 13  
US-10-173-706-574  
; Sequence 574, Application US/10173706  
; Publication No. US20030022293A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C7  
; CURRENT APPLICATION NUMBER: US/10/173,706  
; CURRENT FILING DATE: 2002-06-17  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-173-706-574

Query Match 54.7%; Score 35; DB 9; Length 882;

QY 1 CFOLQNRMKV 11  
Db 848 CEQVDRNIRRV 858

RESULT 10  
US-10-174-590-574  
; Sequence 574, Application US/10174590  
; Publication No. US20030008352A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C42  
; CURRENT APPLICATION NUMBER: US/10/174,590  
; CURRENT FILING DATE: 2002-06-18  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-174-590-574

Query Match 54.7%; Score 35; DB 9; Length 882;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLQNRMKV 11  
Db 848 CEQVDRNIRRV 858

RESULT 11  
US-10-176-758-574  
; Sequence 574, Application US/10176758  
; Publication No. US20030008353A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C104  
; CURRENT APPLICATION NUMBER: US/10/176,758  
; CURRENT FILING DATE: 2002-06-21  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-758-574

Query Match 54.7%; Score 35; DB 9; Length 882;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQLQRMNRKV 11  
| | | | | | | | | |  
Db 848 CEQVDRNIRRV 858

## RESULT 14

US-10-175-738-574  
; Sequence 574, Application US/10175738  
; Publication No. US20030022294A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430RIC45  
; CURRENT APPLICATION NUMBER: US/10/175,738  
; CURRENT FILING DATE: 2002-06-19  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-738-574

Query Match 54.7%; Score 35; DB 9; Length 882;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQLQRMNRKV 11  
| | | | | | | | | |  
Db 848 CEQVDRNIRRV 858

## RESULT 15

US-10-175-752-574  
; Sequence 574, Application US/10175752  
; Publication No. US20030022295A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430RIC60  
; CURRENT APPLICATION NUMBER: US/10/175,752  
; CURRENT FILING DATE: 2002-06-19  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 574  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-175-752-574

Query Match 54.7%; Score 35; DB 9; Length 882;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQLQRMNRKV 11  
| | | | | | | | | |  
Db 848 CEQVDRNIRRV 858

Search completed: February 21, 2003, 08:08:07  
Job time : 10.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-80  
Perfect score: 64  
Sequence: 1 CFOLQRMKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 73: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	90.6	711	1 TFHUL	lactotransferrin p
2	41	64.1	223	3 T37974	probable peroxisom
3	41	64.1	238	2 T40568	hypothetical prote
4	41	64.1	708	2 JC2323	lactoferrin - goat
5	40	62.5	887	2 H96515	hypothetical prote
6	38	59.4	33	2 S52107	lactoferrin - shee
7	38	59.4	335	2 T33211	hypothetical prote
8	37	57.8	187	2 AH1866	hypothetical prote
9	37	57.8	252	2 F65084	hypothetical prote
10	37	57.8	346	2 T19980	hypothetical prote
11	37	57.8	887	2 A96516	hypothetical prote
12	36	56.2	153	2 A97524	hypothetical prote
13	36	56.2	178	2 AB2743	hypothetical prote
14	36	56.2	228	2 G64908	probable transcrip
15	36	56.2	228	2 B85720	hypothetical prote
16	36	56.2	228	2 E30897	hypothetical prote
17	36	56.2	249	2 AI3401	arginyltransferase
18	36	56.2	389	2 T03612	chalcone synthase
19	36	56.2	409	2 S47440	secy protein - red
20	36	56.2	508	2 AC2342	hypothetical prote
21	36	56.2	519	2 T31292	glycolate oxidase
22	36	56.2	664	2 T28852	probable potassium
23	36	56.2	696	2 AE1210	teichoic acid bios
24	36	56.2	921	2 S49965	probable membrane
25	36	56.2	1166	2 A39432	ATP-dependent deox
26	35	54.7	47	2 T29970	hypothetical prote
27	35	54.7	57	2 D81949	hypothetical prote
28	35	54.7	81	2 E95172	hypothetical prote
29	35	54.7	81	2 E38038	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N;Alternate names: lactoferrin

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C;Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R;Cho, Y.

submitted to the EMBL Data Library, March 1994

A;Reference number: G06820

A;Accession: G01394

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-711 <CHO>

A;Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R;Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5286, 1990

A;Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A;Reference number: S11228; MUID:90384839; PMID:2402455

A;Accession: S11228

A;Molecule type: mRNA

A;Residues: 1-148,'T',150-422,'C',424-711 <REY>

A;Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R;Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A;Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A;Reference number: A45401; MUID:93125571; PMID:1480183

A;Accession: A45401

A;Molecule type: DNA

A;Residues: 1-15 <TEN>

A;Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A;Experimental source: Placenta

A;Note: sequence extracted from NCBI backbone (NCBIP:122202)

R;Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A;Title: Nucleotide sequence of human lactoferrin cDNA.

A;Reference number: S10324; MUID:90326549; PMID:2374734

A;Accession: S10324

A;Molecule type: mRNA

A;Residues: 3-711 <POW>

A;Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R;Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A;Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A;Reference number: S15853; MUID:91264786; PMID:2049066

A;Accession: S15853

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 20-31 <ST1>

A;Accession: S20841

A;Molecule type: protein

A;Residues: 20-28,'X',30-31 <ST2>

hypothetical prote  
hypothetical prote  
60S ribosomal prot  
ribosomal protein  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable malate de  
malate oxidoreduct  
hypothetical prote  
embryogenic callus  
probable ribonucle  
hypothetical prote  
2-amino-4-hydroxy-  
protein ZK1073.2 [  
hypothetical prote

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:g186815; PIDN:AAA86665.1; PID:g386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SWKPEVNV' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-400  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norrskov, L.  
 Eur. J. Biochem. 241, 303-306, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:ITF  
 A;Cross-references: GDB:119369; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status experimental  
 F;157,498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-534, 513-527, 524-537, 647-652/Disulfide bonds: #status experimental

Query Match 90.6%; Score 58; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0053;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNMKVR 12  
 ||| |||||  
 Db 39 CFQWQNMKVR 50

RESULT 2  
 T37974  
 probable peroxisomal membrane protein - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
 A;Accession: T37974  
 R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, March 1996  
 A;Reference number: Z21759  
 A;Accession: T37974  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-223 <MUR>  
 A;Cross-references: EMBL:Z69909; PIDN:CAA93785.1; GSPDB:GN00066; SPDB:SPAC19G10.03C  
 A;Experimental source: strain 972h-; cosmid c19G10  
 C;Genetics:  
 A;Gene: SPDB:SPAC19G10.03C

A;Map position: 1  
 A;Introns: 10/3; 170/2  
 C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC582.09  
 C;Keywords: peroxisome

Query Match 64.1%; Score 41; DB 2; Length 223;  
 Best Local Similarity 41.7%; Pred. No. 2.9;  
 Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNMKVR 12  
 ||| |||||  
 Db 140 CYEQQNSKKIK 151

## RESULT 3

T40568

hypothetical protein SPBC582.09 - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000

C;Accession: T40568

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.

submitted to the EMBL Data Library, July 1999

A;Reference number: Z21937

A;Accession: T40568

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-238 &lt;WOO&gt;

A;Cross-references: EMBL:AL096788; PIDN:CA846672.1; GSPDB:GN00067; SPDB:SPBC582.09

A;Experimental source: strain 972h-; cosmid c582

C;Genetics:

A;Gene: SPDB:SPBC582.09

A;Map position: 2

A;Introns: 15/3; 25/3; 185/2

C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC582.09

Query Match 64.1%; Score 41; DB 2; Length 238;  
 Best Local Similarity 41.7%; Pred. No. 3.1;  
 Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLOQNMKVR 12  
 ||| |||||  
 Db 155 CYEQQNSKKIK 166

## RESULT 4

JC2323

lactoferrin - goat

C;Species: Capra aegagrus hircus (domestic goat)

C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999

C;Accession: JC2323

R;Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.

Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994

A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant loci

A;Reference number: JC2323; MUID:94380047; PMID:8093048

A;Accession: JC2323

A;Molecule type: mRNA

A;Residues: 1-708 &lt;LEP&gt;

A;Superfamily: transferrin; transferrin repeat homology

C;Keywords: duplication; glycoprotein

F;252,300/Domain: transferrin repeat homology <TRH2>  
 F;252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.1%; Score 41; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 9.1;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLOQNMKVR 11  
 ||| |||||  
 Db 38 CYQWQNMKRL 48

## RESULT 5

H96515

Qy	1	CFQLQR--NMKKVR	12
		:	:
Db	106	CFSLRMANLRKQV	119



A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: ydfH

C;Superfamily: hypothetical protein b1540

C;Keywords: DNA binding; transcription regulation

Query Match 56.2%; Score 36; DB 2; Length 228;

Best Local Similarity 41.7%; Pred. No. 27;

Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLQRMKVR 12

|:|:|:|:

DB 110 CYQLQNHLQQR 121

#### RESULT 15

B85720

Hypothetical protein ydfH [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 01-Mar-2002

C;Accession: B85720

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: B85720

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-228 <STO>

A;Cross-references: GB:AE005174; NID:g12515116; PIDN:AAG56222.1; GSPDB:GN00145; UWGP:421

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: ydfH

C;Superfamily: hypothetical protein b1540

Query Match

56.2%; Score 36; DB 2; Length 228;

Best Local Similarity 41.7%; Pred. No. 27;

Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLQRMKVR 12

|:|:|:|:

DB 110 CYQLQNHLQQR 121

Search completed: February 21, 2003, 07:47:53

Job time : 11.65 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-80  
Perfect score: 64  
Sequence: 1 CFQLQNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	58	90.6	711	1 TRFL_HUMAN	P02788 homo sapien
2	45	70.3	695	1 TRFL_HORSE	O77811 equus caball
3	41	64.1	238	1 YBM9_SCHPO	Q10333 schizosacch
4	41	64.1	708	1 TRFL_CAMDR	Q9tumo0 camelus dro
5	41	64.1	708	1 TRFL_CAPHI	Q29477 capra hircu
6	38	59.4	573	1 UREL_LACPE	P26929 lactobacill
7	37	57.8	226	1 NUKM_NEUCR	O47950 neurospora
8	37	57.8	252	1 YGHR_ECOLI	Q46842 escherichia
9	37	57.8	252	1 FRQ_LEPAU	Q01115 leptospheer
10	36	56.2	228	1 YDPH_ECOLI	P77577 escherichia
11	36	56.2	410	1 SECY_CVACA	P46249 cyanidium c
12	36	56.2	921	1 YIB7_YEAST	P40547 saccharomyc
13	36	56.2	1166	1 ADBE_BACSU	P23477 bacillus su
14	36	56.2	3674	1 SPOR_HUMAN	Q9nrc66 homo sapien
15	35	54.7	179	1 RK27_TOBAC	P30155 nicotiana t
16	35	54.7	234	1 YX02_CAEEL	Q11109 caenorhabdi
17	35	54.7	265	1 Y125_MYCCA	P53661 mycoplasma
18	35	54.7	795	1 DEGY_CAEEL	O01635 caenorhabdi
19	34	53.1	864	1 STAZ_PIG	O02799 sus scrofa
20	34	53.1	3678	1 DMD_MOUSE	P11531 mus musculu
21	34	53.1	3680	1 DMD_CANFA	O97592 canis fami
22	34	53.1	3685	1 DMD_HUMAN	P11532 homo sapien
23	33	51.6	192	1 RL24_SCHPO	Q10353 schizosacch
24	33	51.6	261	1 UL31_HVSA	Q01041 herpesvirus
25	33	51.6	309	1 LTRA_KLEPN	P52689 klebsiella
26	33	51.6	410	1 IIBC_FELCA	Q9mzv66 felis silve
27	33	51.6	440	1 PCXA_SYNY3	P75028 synchocyst
28	33	51.6	510	1 YAGF_SCHPO	Q09877 schizosacch
29	33	51.6	554	1 RECN_RICPR	Q9zdy2 rickettsia
30	33	51.6	604	1 LPXE_CHLPN	Q9z6u3 chlamydia p
31	33	51.6	604	1 SR68_DROME	Q9vss2 drosophila
32	33	51.6	656	1 YC26_PORPU	P51392 porphyra pu
33	33	51.6	770	1 MAO1_RHIME	O30807 rhizobium m

## RESULT 1

ID	TRFL_HUMAN	STANDARD;	PRT;	711 AA.
AC	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;			
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;			
DE	Lactoferrin B; Lactoferrin C].			
GN	LTF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI TaxID=9606;			
RN	[1] _			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cho Y.Y.;			
RL	Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	Connely O.N.;			
RL	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RT	"Molecular cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RL	sequences.";			
RN	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RT	"cDNA cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8]			

34	33	51.6	842	1 ORP7_HUMAN	Q9bzf2 homo sapien
35	33	51.6	3164	1 TEGU_HSV11	P10220 herpes simp
36	32.5	50.8	358	1 OAS1_RAT	Q05961 rattus norv
37	32.5	50.8	367	1 OASA_MOUSE	P11928 mus musculu
38	32	50.0	115	1 YB92_BALSO	Q8y052 raietonia s
39	32	50.0	148	1 ANGI_BOVIN	P10152 bos taurus
40	32	50.0	177	1 YAI6_SCHPO	Q09688 schizosacch
41	32	50.0	195	1 AANT_HVDV3	Q29996 hepatitis d
42	32	50.0	198	1 RK27_ARATH	Q9fln4 arabidopsis
43	32	50.0	245	1 YD76_METUA	Q58771 methanococc
44	32	50.0	267	1 RS3A_DROME	P55830 drosophila
45	32	50.0	334	1 ANM4_HUMAN	Q9nr22 homo sapien

## ALIGNMENTS



RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogdan J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RT Legrand D., Spik G., Montreuil J., Jolles P.;  
RL "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RT Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RT Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dregan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RT Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253--methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RN Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -!- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53961; CRA37914.1; -  
DR EMBL; U07643; AAB60324.1; -  
DR EMBL; M93150; AAA36159.1; -  
DR EMBL; M83202; AAA59511.1; -  
DR EMBL; M83205; AAA58656.1; -  
DR EMBL; M18642; AAA66665.1; -  
DR EMBL; AF332168; AAG48753.1; -  
DR EMBL; BC015823; AAH15822.1; -  
DR EMBL; BC015823; AAH15823.1; -  
DR EMBL; M73700; AAA59479.1; -  
DR EMBL; X52941; CAA37116.1; -  
DR EMBL; U95626; AAB57795.1; -  
DR PIR; S11228; TFHUL.  
DR PDB; 1LFC; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LFH; 31-OCT-93.  
DR PDB; 1LFI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1LKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

Query Match 90.6%; Score 58; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred.No. 0.0014;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLQNNRKVR 12  
 ||| |||||  
 Db 39 CFQWRNNRKVR 50

## RESULT 2

TRFL HORSE  
 ID TRFL HORSE STANDARD; PRT; 695 AA.  
 AC 077811;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactoferrin precursor (Lactoferrin) (Fragment).  
 GN LTF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 [1]  
 SEQUENCE FROM N.A.  
 RP Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RA "cDNA sequence of mare lactoferrin."  
 RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RC TISSUE=Milk;  
 RX MEDLINE=99296631; PubMed=10366507;  
 RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "Three-dimensional structure of mare lactoferrin at 2.6-A  
 resolution."  
 RL J. Mol. Biol. 289:303-317(1999).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AJ010930; CAA09407.1; -.  
 DR PDB; 1B1X; 02-DEC-98.  
 DR PDB; 1B7U; 02-FEB-99.  
 DR PDB; 1B7Z; 02-FEB-99.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 1.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal; 3D-structure.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 6  
 FT CHAIN 7 695 LACTOTRANSFERRIN.  
 FT REPEAT 7 350 1.  
 FT REPEAT 351 695 2.  
 FT DISULFID 15 51  
 FT DISULFID 25 42  
 FT DISULFID 121 204

FT DISULFID 163 179  
 FT DISULFID 166 189  
 FT DISULFID 176 187  
 FT DISULFID 237 251  
 FT DISULFID 354 386  
 FT DISULFID 364 377  
 FT DISULFID 411 690  
 FT DISULFID 431 653  
 FT DISULFID 463 538  
 FT DISULFID 487 681  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66 66  
 FT METAL 98 98  
 FT METAL 198 198  
 FT METAL 259 259  
 FT METAL 401 401  
 FT METAL 439 439  
 FT METAL 532 532  
 FT METAL 601 601  
 FT BINDING 127 127  
 FT BINDING 167 167  
 FT BINDING 143 143  
 FT CARBOHYD 287 287  
 FT CARBOHYD 482 482  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 70.3%; Score 45; DB 1; Length 695;  
 Best Local Similarity 66.7%; Pred.No. 0.45;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQLQNNRKVR 12  
 | : ||| : |||  
 Db 25 CAFQNNRKVR 36

## RESULT 3

YEM9 SCHPO  
 ID YEM9 SCHPO STANDARD; PRT; 238 AA.  
 AC Q10333;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C582.09 in chromosome II.  
 GN SPBC582.09.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21849401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins S., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert B., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Calibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armetrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
RT "The genome sequence of *Schizosaccharomyces pombe*."  
RL Nature 415:871-880(2002).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AL096788; CAB46672.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 238 AA; 26479 MW; 58095AA8CD708180 CRC64; -  
Query Match 64.1%; Score 41; DB 1; Length 238;  
Best Local Similarity 41.7%; Pred. No. 0.85;  
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CFQLQRMKVR 12  
|:|:|:|:|:  
Db 155 CYELQNSKKIK 166  
-----  
RESULT 4  
TRFL\_CAMDR STANDARD; PRT; 708 AA.  
ID Q9TUM0; Q9MZS5;  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactoferrin precursor (Lactoferrin).  
GN LTF.  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
OX NCBI\_TaxID=9838;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin."  
RL Int. Dairy J. 9:481-486(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AJ131674; CAB53387.1; -.  
CC EMBL; AF165879; AAF82241.1; -.

DR HSP; O77811; 1B1X.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 2.  
DR PRINTS; PR00422; TRANSFERRIN.  
DR SMART; SM00094; TR\_FER; 2.  
DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
KW Signal.  
FT SIGNAL 1 19 BY SIMILARITY.  
FT CHAIN 20 708 LACTOTRANSFERRIN.  
FT REPEAT 20 363 1.  
FT REPEAT 364 708 2.  
FT DISULFID 28 64 BY SIMILARITY.  
FT DISULFID 38 55 BY SIMILARITY.  
FT DISULFID 134 217 BY SIMILARITY.  
FT DISULFID 176 192 BY SIMILARITY.  
FT DISULFID 189 200 BY SIMILARITY.  
FT DISULFID 250 264 BY SIMILARITY.  
FT DISULFID 367 399 BY SIMILARITY.  
FT DISULFID 377 390 BY SIMILARITY.  
FT DISULFID 424 703 BY SIMILARITY.  
FT DISULFID 444 666 BY SIMILARITY.  
FT DISULFID 476 551 BY SIMILARITY.  
FT DISULFID 500 694 BY SIMILARITY.  
FT DISULFID 510 524 BY SIMILARITY.  
FT DISULFID 521 534 BY SIMILARITY.  
FT DISULFID 592 606 BY SIMILARITY.  
FT DISULFID 644 649 BY SIMILARITY.  
FT METAL 79 79 IRON 1 (BY SIMILARITY).  
FT METAL 111 111 IRON 1 (BY SIMILARITY).  
FT METAL 211 211 IRON 1 (BY SIMILARITY).  
FT METAL 272 272 IRON 1 (BY SIMILARITY).  
FT METAL 414 414 IRON 2 (BY SIMILARITY).  
FT METAL 452 452 IRON 2 (BY SIMILARITY).  
FT METAL 545 545 IRON 2 (BY SIMILARITY).  
FT METAL 614 614 IRON 2 (BY SIMILARITY).  
FT BINDING 140 140 ANION (BY SIMILARITY).  
FT BINDING 482 482 ANION (BY SIMILARITY).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 261 261 F -> S (IN REF. 2).  
FT CONFLICT 304 304 G -> A (IN REF. 2).  
FT CONFLICT 330 330 S -> P (IN REF. 2).  
FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
FT CONFLICT 506 506 L -> F (IN REF. 2).  
FT CONFLICT 609 609 A -> P (IN REF. 2).  
FT CONFLICT 642 642 R -> Q (IN REF. 2).  
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;  
Query Match 64.1%; Score 41; DB 1; Length 708;  
Best Local Similarity 66.7%; Pred. No. 2.7;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 CFQLQRMKVR 12  
|:|:|:|:|:  
Db 38 CAQWQRMKVR 49  
-----  
RESULT 5  
TRFL\_CAMDR STANDARD; PRT; 708 AA.  
ID Q9TUM0; Q9MZS5;  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactoferrin precursor (Lactoferrin).  
GN LTF.  
OS Capra hircus (Goat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 56 56 I -> V (IN REF. 2).
FT CONFLICT 88 88 L -> R (IN REF. 2).
FT CONFLICT 124 124 Q -> K (IN REF. 2).
FT CONFLICT 154 154 F -> P (IN REF. 2).
FT CONFLICT 304 304 S -> R (IN REF. 2).
FT CONFLICT 414 414 D -> G (IN REF. 2).
SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;

Query Match 64.1%; Score 41; DB 1; Length 708;
Best Local Similarity 63.6%; Pred. No. 2.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLQRNMKRV 11
Db 38 CYQWRMRMKL 48
.:|:||||:

RESULT 6
UREI_LACFE STANDARD; PRT; 573 AA.
ID UREI_LACFE AC P26929;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acid urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN UREC.
OS Lactobacillus fermentum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1613;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=JCM 5869;
RA Suzuki K., Takahashi M., Imamura S., Ishikawa T.;
RL Submitted (XXXX-1992) to the EMBL/GenBank/DDBJ databases.
CC -|- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -|- COPACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -|- SUBUNIT: (ALPHA, BETA, GAMMA) (3) (BY SIMILARITY).
CC -|- FM: Iys-222 is carbamylated. The carbamoyl group provides the ligands for the two nickel ions (By similarity).
CC -|- SIMILARITY: BELONGS TO THE UREASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
DR EMBL; D10605; BAA01460.1; --
DR HSP; P41020; LUEP.
DR MKRPS; M38.UNW; -.
DR InterPro; IPR001924; UreaseA.
DR Pfam; PF00449; urease; 1.
DR Pfam; PF02802; urease_C; 1.
DR ProDom; PD002467; UreaseA; 1.
DR PROSITE; PS00145; UREASE_2; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR Hydrolase; Metal-binding; Nickel.
KM METAL 139 139 NICKEL 2 (BY SIMILARITY).
FT METAL 141 141 NICKEL 2 (BY SIMILARITY).
FT METAL 222 222 NICKEL 1 AND 2 (BY SIMILARITY).
FT METAL 251 251 NICKEL 1 (BY SIMILARITY).
FT METAL 277 277 NICKEL 1 (BY SIMILARITY).
FT METAL 365 365 NICKEL 2 (BY SIMILARITY).
FT ACT SITE 325 325 BY SIMILARITY.

```

```

SQ .SEQUENCE 573 AA; 61805 MW; 2D2619781C39E54B CRC64;
Query Match 59.4%; Score 38; DB 1; Length 573;
Best Local Similarity 58.3%; Pred. No. 8.4;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLQRMNRKVR 12
|:|:|:|:|:|
Db 379 CQQLASRMKVR 390

RESULT 7
NUKM NEUCR STANDARD; PRT; 226 AA.
AC A047950;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 19.3 kDa subunit, mitochondrial
DE precursor (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-19.3KD) (CI-19.3KD).
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN SEQUENCE FROM N.A.
RP MEDLINE=92233617; PubMed=10216160;
RX Sousa R., Barquera B., Duarte M., Pinel M., Videira A.;
RT "Characterisation of the last Fe-S cluster-binding subunit
of Neurospora crassa complex I.";
RL Biochim. Biophys. Acta 1411:142-146(1999).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- COFACTOR: MAY CONTAIN AN IRON-SULFUR 4FE-4S CLUSTER.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 20 kDa SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC EMBL; AJ001520; CAA04802.1; -
DR InterPro; IPR002096; Cmp1xi_20kDa.
DR Pfam; PF01059; oxidored g6; 1.
DR PROSITE; PS01150; COMPLEX1_20K; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transit peptide;
KW Iron-sulfur; 4Fe-4S. ?
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN 1 ? 226 NADH-UBIQUINONE OXIDOREDUCTASE 19.3 KDA
SUBUNIT.
FT METAL 101 101 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 102 102 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 166 166 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 196 196 IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ SEQUENCE 226 AA; 24972 MW; E48CB09C64AAAF6C CRC64;

Query Match 57.8%; Score 37; DB 1; Length 226;
Best Local Similarity 63.6%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQLQRMNRKVR 12
|:|:|:|:|:|
Db 208 FQLQRMNRKTK 218

RESULT 8
YGHR ECOLI STANDARD; PRT; 252 AA.
ID YGHR ECOLI
SQ .SEQUENCE 573 AA; 61805 MW; 2D2619781C39E54B CRC64;
Query Match 59.4%; Score 38; DB 1; Length 573;
Best Local Similarity 58.3%; Pred. No. 8.4;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLQRMNRKVR 12
|:|:|:|:|:|
Db 379 CQQLASRMKVR 390

RESULT 7
NUKM NEUCR STANDARD; PRT; 226 AA.
AC A047950;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 19.3 kDa subunit, mitochondrial
DE precursor (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-19.3KD) (CI-19.3KD).
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN SEQUENCE FROM N.A.
RP MEDLINE=92233617; PubMed=10216160;
RX Sousa R., Barquera B., Duarte M., Pinel M., Videira A.;
RT "Characterisation of the last Fe-S cluster-binding subunit
of Neurospora crassa complex I.";
RL Biochim. Biophys. Acta 1411:142-146(1999).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- COFACTOR: MAY CONTAIN AN IRON-SULFUR 4FE-4S CLUSTER.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 20 kDa SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC EMBL; U28377; AAA69151.1; -
DR EMBL; AB000381; AAC76020.1; -
DR EcoGene; EGI3002; YGHR.
KW Hypothetical protein; ATP-binding; Complete proteome.
FT NP BIND 28 35 ATP (POTENTIAL).
SQ SEQUENCE 252 AA; 28151 MW; E9C49E970425C758 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 252;
Best Local Similarity 57.1%; Pred. No. 5.4;
Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 CFQLQRMNRKVR 12
|:|:|:|:|:|
Db 106 CFSLRMANLRKVV 119

RESULT 9
FRQ LEPAU STANDARD; PRT; 992 AA.
ID FRQ LEPAU
AC Q01115;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Frequency clock protein.
FN Frequency clock protein.
GN FRQ.
OS Leptospira australiensis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Leptosphaeriaceae; Leptosphaeria.
OX NCBI_TaxID=40116;
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 66919; PubMed=9037100;
RX MEDLINE=97188515;
RA Lewis M.T., Morgan L.W., Feldman J.F.;
RT "Analysis of frequency (frq) clock gene homologs: evidence for a
helix-turn-helix transcription factor.";
RL Mol. Gen. Genet. 253:401-414(1997).
CC -!- FUNCTION: CIRCADIAN CLOCK COMPONENT INVOLVED IN THE GENERATION OF
BIOLOGICAL RHYTHMS, IN PARTICULAR IN RHYTHM STABILITY, PERIOD
LENGTH, AND TEMPERATURE COMPENSATION. BEHAVES AS A NEGATIVE
ELEMENT IN CIRCADIAN TRANSCRIPTIONAL LOOP (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U25851; AAB96844.1; --  
 CC Biological rhythms; Transcription regulation; Nuclear protein.  
 CC DOMAIN 550 554 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC DOMAIN 868 890 ASP/GLU-RICH (ACIDIC).  
 CC SEQUENCE 992 AA; 107559 MW; 98E029B585F83AFF CRC64;

Query Match 57.88; Score 37; DB 1; Length 992;  
 Best Local Similarity 72.78; Pred. No. 24;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQLQNRMRKVR 12  
 |||||:||||  
 Db 504 FQLSRDGRKVR 514

## RESULT 10

YDFH\_ECOLI STANDARD; PRT; 228 AA.  
 AC P77577;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical transcriptional regulator ydfH.  
 GN YDFH OR B1540 OR Z2157 OR EC52149.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 CC NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "the complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kasai H., Kimura S., Kitakawa M.,  
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
 RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,  
 RA Ohshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,  
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377 (1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11208551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potancusis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533 (2001).

[4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:111-22 (2001).  
 CC -|- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query Match 56.2%; Score 36; DB 1; Length 228;  
 Best Local Similarity 41.7%; Pred. No. 7.6;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLQNRMRKVR 12  
 |||||:||||  
 Db 110 CYCLEQLHQOR 121

## RESULT 11

SECV\_CYACA STANDARD; PRT; 410 AA.  
 AC P46249; Q9MD55;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 18-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Preprotein translocase secY subunit.  
 GN SECY.  
 OS Cyanidium caldarium.  
 OG Chloroplast  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;  
 OC Cyanidium.  
 CC NCBI\_TaxID=2771;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RK-1;  
 RX MEDLINE=97134960; PubMed=8980520;  
 RA Vogel H., Fischer S., Valentin K.-U.;  
 RT "A model for the evolution of the plastid sec apparatus inferred from  
 RT secY gene phylogeny.";  
 RL Plant Mol. Biol. 32:685-692 (1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RK-1;  
 RX MEDLINE=20496959; PubMed=11040290;



CC ENDONUCLEASE, ATP-DEPENDENT HELICASE AND DNA-DEPENDENT ATPASE  
CC ACTIVITIES  
CC -!- SUBUNIT: THE B.SUBTILIS ATP-DEPENDENT NUCLEASE COMPLEX IS FORMED  
CC BY ONE B SUBUNIT IN CONJUNCTION WITH ONE A SUBUNIT.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; M63489; AAA22200.1; -  
CC EMBL; Y14081; CAA74481.1; -  
CC EMBL; Z39109; CAB12902.1; -  
CC PIR; A39432; A39432.  
CC Subtilisin; BG10465; addB.  
CC InterPro; IPR000212; UvrD-helicase; 1.  
CC Pfam; PF00580; UvrD-helicase; 1.  
CC Hydrolase; Nuclease; Exonuclease; Helicase; ATP-binding;  
KW DNA repair; Complete proteome.  
FT NP\_BIND 1 22 ATP (POTENTIAL).  
SQ SEQUENCE 1166 AA; 134631 MW; 1A57BBE81A08AB4A CRC64;  
Query Match 56.2%; Score 36; DB 1; Length 1166;  
Best Local Similarity 54.5%; Pred. No. 44;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 2 FOLQRMKVR 12  
Db 486 FOLQRMKAK 496  
RESULT 14  
ID SPCR HUMAN STANDARD; PRT; 3674 AA.  
AC Q9NRG6;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Spectrin beta chain, brain 4 (Spectrin, non-erythroid beta chain 4)  
DE (Beta-v spectrin) (BSPECV).  
GN SPTBN5 OR SPTBN4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Retina, Cerebellum, and Spinal cord;  
RX MEDLINE=20347255; PubMed=10764729;  
RA Stabach P.R.; Morrow J.S.;  
RT "Identification and characterization of beta V spectrin, a mammalian  
RT ortholog of Drosophila beta H spectrin.";  
RL J. Biol. Chem. 275:21385-21395(2000).  
CC -!- SUBUNIT: Probably associates with an alpha chain.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Detected prominently in the  
CC outer segments of photoreceptor rods and cones and in the  
CC basolateral membrane and cytosol of gastric epithelial cells.  
CC -!- TISSUE SPECIFICITY: Expressed at very low levels in many tissues,  
CC with strongest expression in cerebellum, spinal cord, stomach,  
CC pituitary gland, liver, pancreas, salivary gland, kidney, bladder,  
CC and heart.  
CC -!- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.  
CC -!- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -!- SIMILARITY: CONTAINS 31 SPECTRIN REPEATS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AF233523; AAF65317.1; -  
CC HSSP; Q01082; 1BKR.  
CC Genew; HGNC:15680; SPTBN5.  
CC MIM; 605916; -  
CC InterPro; IPR001589; Actbind actnin.  
CC InterPro; IPR001715; Calponin-like.  
CC InterPro; IPR001849; PH.  
CC InterPro; IPR002017; Spectrin.  
CC InterPro; IPR001605; Spectrin\_PH.  
CC Pfam; PF00169; PH; 1.  
CC Pfam; PF00307; CH; 2.  
CC Pfam; PF00435; spectrin; 30.  
CC PRINTS; P000683; SPECTRINPH.  
CC SMART; SM00033; CH; 2.  
CC SMART; SM00233; PH; 1.  
CC SMART; SM00150; SPEC; 29.  
CC PROSITE; PS00019; ACTININ\_1; 1.  
CC PROSITE; PS00020; ACTININ\_2; FALSE\_NEG.  
CC PROSITE; PS50021; CH; 2.  
CC PROSITE; PS50003; PH\_DOMAIN; 1.  
KW Cytoskeleton; Membrane; Repeat; Actin-binding; Capping protein.  
FT DOMAIN 1 279 ACTIN-BINDING (BY SIMILARITY).  
FT DOMAIN 54 159 CH 1.  
FT DOMAIN 177 279 CH 2.  
FT REPEAT 306 416 SPECTRIN 1.  
FT REPEAT 426 531 SPECTRIN 2.  
FT REPEAT 533 637 SPECTRIN 3.  
FT REPEAT 639 743 SPECTRIN 4.  
FT REPEAT 745 810 SPECTRIN 5.  
FT REPEAT 867 894 SPECTRIN 6.  
FT REPEAT 896 998 SPECTRIN 7.  
FT REPEAT 1102 1206 SPECTRIN 8.  
FT REPEAT 1208 1312 SPECTRIN 9.  
FT REPEAT 1314 1417 SPECTRIN 10.  
FT REPEAT 1419 1488 SPECTRIN 11.  
FT REPEAT 1520 1624 SPECTRIN 12.  
FT REPEAT 1626 1728 SPECTRIN 13.  
FT REPEAT 1730 1835 SPECTRIN 14.  
FT REPEAT 1837 1941 SPECTRIN 15.  
FT REPEAT 1943 2047 SPECTRIN 16.  
FT REPEAT 2049 2147 SPECTRIN 17.  
FT REPEAT 2149 2253 SPECTRIN 18.  
FT REPEAT 2255 2301 SPECTRIN 19.  
FT REPEAT 2314 2362 SPECTRIN 20.  
FT REPEAT 2364 2468 SPECTRIN 21.  
FT REPEAT 2470 2574 SPECTRIN 22.  
FT REPEAT 2576 2680 SPECTRIN 23.  
FT REPEAT 2682 2785 SPECTRIN 24.  
FT REPEAT 2787 2891 SPECTRIN 25.  
FT REPEAT 2893 2997 SPECTRIN 26.  
FT REPEAT 2999 3103 SPECTRIN 27.  
FT REPEAT 3105 3210 SPECTRIN 28.  
FT REPEAT 3212 3312 SPECTRIN 29.  
FT REPEAT 3314 3418 SPECTRIN 30.  
FT REPEAT 3420 3482 SPECTRIN 31.  
FT DOMAIN 3533 3641 PH.  
SQ SEQUENCE 3674 AA; 416832 MW; 8C0FDFADDD0C7C18 CRC64;  
Query Match 56.2%; Score 36; DB 1; Length 3674;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CFOLQRMKVR 12  
Db 2290 CLQRRRLREFR 2301  
RESULT 15  
RK27\_TOBAC



```

ID  RX27_TOBAC  STANDARD;  PRT;  179 AA.
AC  P30155;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  01-FEB-1996 (Rel. 33, Last annotation update)
DE  50S ribosomal protein L27, chloroplast precursor (CL27).
GN  RPL27 OR RPL27A.
OS  Nicotiana tabacum (Common tobacco).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX  NCBI_TaxID=4097;
RN  [1]
RP  SEQUENCE FROM N.A., AND SEQUENCE OF 52-68.
RC  STRAIN=cv. Petit Havana; TISSUE=Leaf;
RX  MEDLINE=92345244; PubMed=1339289;
RA  Elhag G.A., Bourque D.P.;
RT  "Nuclear-encoded chloroplast ribosomal protein L27 of Nicotiana
RL  tabacum: cDNA sequence and analysis of mRNA and genes.";
RL  Biochemistry 31:6856-6864(1992)
CC  -!- SIMILARITY: BELONGS TO THE L27P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M75731; AAA34115.1; -.
DR  EMBL; M98473; AAA34104.1; -.
DR  PIR; A42840; A42840.
DR  InterPro; IPR001684; Ribosomal_L27.
DR  Pfam; PF01016; Ribosomal_L27; 1.
DR  PRINTS; PR00063; RIBOSOMALI27.
DR  ProDom; PD003114; Ribosomal_L27; 1.
DR  TIGRams; TIGR00062; L27; 1.
DR  PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW  Ribosomal protein; Chloroplast; Transit peptide.
FT  TRANSIT 1 51 CHLOROPLAST.
FT  CHAIN 52 179 50S RIBOSOMAL PROTEIN L27.
SQ  SEQUENCE 179 AA; 19664 MW; D6A30E2E90CB3EB7 CRC64;

Query Match      54.7%; Score 35; DB 1; Length 179;
Best Local Similarity 63.6%; Pred.No. 9.2;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  2 FOLQNNMKVR 12
Db  152 FQLQRRKAR 162

```

Search completed: February 21, 2003, 07:27:55  
 Job time : 5.6 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds

(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107b-80  
Perfect score: 64  
Sequence: 1 CFQLQRNRKRV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL 21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	53	82.8	711	4	Q8TCD2
2	49	76.6	38	4	Q9UCY5
3	40	62.5	253	5	Q9VZQ6
4	40	62.5	315	5	Q95SH7
5	40	62.5	887	10	Q9SX85
6	39.5	61.7	735	10	Q9LGI7
7	38	59.4	33	6	Q9TR80
8	38	59.4	125	4	Q9H960
9	38	59.4	335	5	O61888
10	38	59.4	573	2	Q9S6F6
11	37	57.8	187	16	Q8YZH9
12	37	57.8	210	3	Q9UW7
13	37	57.8	346	5	O45308
14	37	57.8	565	16	Q8XJU2
15	37	57.8	887	10	Q9SX86
16	36	56.2	121	10	Q9L7N4

17	36	56.2	178	16	Q8UEP3
18	36	56.2	249	16	Q8YGG0
19	36	56.2	389	10	Q43595
20	36	56.2	488	5	Q9VFF6
21	36	56.2	508	16	Q8YPA5
22	36	56.2	519	2	O86000
23	36	56.2	528	5	Q9N9X2
24	36	56.2	675	5	Q9GYM8
25	36	56.2	696	16	Q8Y833
26	36	56.2	3680	5	Q9VR08
27	35	54.7	47	5	Q23578
28	35	54.7	57	16	Q9JV34
29	35	54.7	81	16	Q97PM0
30	35	54.7	104	16	Q97BE1
31	35	54.7	129	3	O59956
32	35	54.7	130	4	Q9H7E7
33	35	54.7	151	16	O84177
34	35	54.7	159	10	O22165
35	35	54.7	192	5	Q9VPD0
36	35	54.7	213	10	Q9AMW7
37	35	54.7	295	10	O65376
38	35	54.7	306	12	O55587
39	35	54.7	426	16	Q9KOD8
40	35	54.7	426	16	Q9JVE6
41	35	54.7	431	3	Q12124
42	35	54.7	445	10	Q9LXD7
43	35	54.7	445	10	Q9FLF7
44	35	54.7	595	10	Q8SLW9
45	35	54.7	644	16	Q9PHN6

#### ALIGNMENTS

##### RESULT 1

ID Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2, 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
EL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -.  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 82.8%; Score 53; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.054;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLQRNRKRV 11

Db 39 CFQWRNRKRV 49

##### RESULT 2

ID Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96081613; PubMed=8551695;  
RA Sato I.;  
RT "Characterization of the 84-kDa protein with ABH activity in human  
RT seminal plasma.";  
RL Jpn. J. Legal Med. 49:281-293 (1995).  
DR HSP; P02788; IBKA.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
DR SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDBEE CRC64;  
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDBEE CRC64;  
Query Match 76.6%; Score 49; DB 4; Length 38;  
Best Local Similarity 90.9%; Pred. No. 0.018;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 FOLQNNMKVR 12  
Db 21 FQWNNMKVR 31  
RESULT 3  
Q9VZQ6 PRELIMINARY; PRT; 253 AA.  
AC Q9VZQ6  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE CG14969 protein.  
DE CG14969  
GN CG14969  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BERKELEY;  
RC STRAIN=BERKELEY;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferrara C., Ferrier S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levisky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs K.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
DR EMBL; AE003477; AAF47763.1; -.  
DR FlyBase; FBgn0035440; CG14969.  
DR SEQUENCE 253 AA; 27767 MW; D50BD3A60E0F0428 CRC64;  
SQ SEQUENCE 253 AA; 27767 MW; D50BD3A60E0F0428 CRC64;  
Query Match 62.5%; Score 40; DB 5; Length 253;  
Best Local Similarity 58.3%; Pred. No. 7;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 CFQLQNNMKVR 12  
Db 105 CFQLQNNMKVR 116  
RESULT 4  
Q95SH7 PRELIMINARY; PRT; 315 AA.  
AC Q95SH7  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE GH26007p.  
DE CG14969.  
GN CG14969.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BERKELEY;  
RC STRAIN=BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY060790; AAL28338.1; -.  
DR FlyBase; FBgn0035440; CG14969.  
DR SEQUENCE 315 AA; 34892 MW; DF66E013854AEA50 CRC64;  
SQ SEQUENCE 315 AA; 34892 MW; DF66E013854AEA50 CRC64;  
Query Match 62.5%; Score 40; DB 5; Length 315;  
Best Local Similarity 58.3%; Pred. No. 8.7;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 CFQLQNNMKVR 12  
Db 167 CFQLQNNMKVR 178  
RESULT 5  
Q9SX85 PRELIMINARY; PRT; 887 AA.  
AC Q9SX85  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE F16N3.18 protein.  
DE F16N3.18  
GN F16N3.18  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. COLOMBIA;  
RC STRAIN=CV. COLOMBIA;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,

RA Lee J., Liu A., Li J., Kremenetskaia I., Gonzalez A., Altafi H.,  
 RA Araujo R., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,  
 RA Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C., Rowley D.,  
 RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,  
 RA Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC F16N3 sequence.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007519; AAD46032.1; -;  
 DR InterPro: IPR002048; EF-hand.  
 DR PROSITE: PS00018; EF\_HAND; UNKNOWN 1.  
 SQ SEQUENCE 887 AA; 100064 MW; 533510830AA5099 CRC64;

Query Match 62.5%; Score 40; DB 10; Length 887;  
 Best Local Similarity 58.3%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQLOQNRKVR 12  
 | : : : : :  
 Db 659 CHQIERNERNVR 670

## RESULT 6

ID Q9LGL7 PRELIMINARY; PRT; 735 AA.  
 AC Q9LGL7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE F14JL6.21.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altafi H., Bei Q., Chin C., Chiu J., Choi E., Conn L.,  
 RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F14JL6 from chromosome  
 I-1";  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiu J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL: AC002304; AAF79335.1; -;  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450; 2.  
 DR PRINTS: PK00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 735 AA; 84661 MW; 4718F6202974F811 CRC64;

Query Match 61.7%; Score 39.5; DB 10; Length 735;  
 Best Local Similarity 66.7%; Pred. No. 25;  
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CFQLOQNRKRV 11  
 | : : : : :  
 Db 165 CYDLOQNRKRI 176

## RESULT 7

ID Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32 (1995).  
 DR HSSP: O77698; 1CE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 59.4%; Score 38; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 2.3;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQLOQNRKRV 11  
 | : : : : :  
 Db 19 CYQWQNRKRL 29

## RESULT 8

ID Q9H960 PRELIMINARY; PRT; 125 AA.  
 AC Q9H960;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE CDNA FLJ12988 fis, clone NT2RP3000080.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;  
 RT "NEDO human cDNA sequencing project";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK023050; BAB14376.1; -;  
 SQ SEQUENCE 125 AA; 13445 MW; 66A1CDD5774E7BD7 CRC64;

Query Match 59.4%; Score 38; DB 4; Length 125;

```

Best Local Similarity 60.0%; Pred. No. 8.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLQRMNRK 10
   |||:|:|
Db 70 CFSLQQNFKK 79

RESULT 9
O61888 PRELIMINARY; PRT; 335 AA.
AC O61888;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 38.7 kDa protein.
GN K10C9.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Du Z., Maggi L.;
RT "The sequence of C. elegans cosmid K10C9.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067944; AAC17673.1; -.
DR InterPro; IPR003002; 7TM chemot.
DR InterPro; IPR000168; 7TM nematode.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF01461; 7tm.4; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 335 AA; 38704 MW; 04A8CEE6CDB8C5F5 CRC64;

Query Match 59.4%; Score 38; DB 5; Length 335;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLQRMNRK 10
   |||:|:|
Db 217 CVQMERNNKK 226

RESULT 10
Q9S6F6 PRELIMINARY; PRT; 573 AA.
AC Q9S6F6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Urease (EC 3.5.1.5) (Urea amidohydrolase).
GN URE.
OS Lactobacillus fermentum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Lactobacillaceae; Lactobacillus.
OC NCBI_TaxID=1613;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=TK1214;
RA Coton E., Visser J.J., Van Vuuren H.J.J.;
RT "Urease operon of Lactobacillus fermentum.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: UREA + H(2)O = CO(2) + 2 NH(3).
CC -|- COFACTOR: EACH CHAIN BINDS TWO NICKEL IONS (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE UREASE FAMILY.
DR EMBL; AF120718; AD22480.1; -.
DR HSP; P41020; IUBP.
DR InterPro; IPR001924; UreaseA.
DR Pfam; PF00449; urease; 1.
DR Pfam; PF02802; urease C; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Hydrolase; Metal-binding; Nickel.
SQ SEQUENCE 573 AA; 61823 MW; 4D325C6C4A21559C CRC64;

Query Match 59.4%; Score 38; DB 2; Length 573;
Best Local Similarity 58.3%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLQRMNRKVR 12
   |||:|:|
Db 379 CWQLASRMKKVR 390

RESULT 11
Q8YZH9 PRELIMINARY; PRT; 187 AA.
AC Q8YZH9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein AL10481.
GN AL10481.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003582; BAB72439.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 187 AA; 21356 MW; 85C13C2EC8831DA9 CRC64;

Query Match 57.8%; Score 37; DB 16; Length 187;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOLQRMNRKVR 12
   |||:|:|
Db 168 FELETNMRKMR 178

RESULT 12
Q9UUT7 PRELIMINARY; PRT; 210 AA.
AC Q9UUT7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Subunit NUKM of protein NADH:ubiquinone oxidoreductase (Complex I)
DE precursor (EC 1.6.99.3).
GN NUKM.

```

OS Yarrowia lipolytica (Candida lipolytica).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Dipodascaceae; Yarrowia.  
 OX NCBI\_taxid=4952;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E 150;  
 RX MEDLINE=20439569; PubMed=10924914;  
 RA Djafarzadeh Andabali R., Kerscher S., Zwicker K., Radermacher M.,  
 RA Lindahl M., Schaegger H., Brandt U.;  
 RT "Biophysical and structural characterization of proton-translocating  
 RT NADH-Dehydrogenase (complex I) from the strictly aerobic yeast  
 RT Yarrowia lipolytica.";  
 RL Biochim. Biophys. Acta 1459:230-238 (2000).  
 DR EMBL: AJ250340; CAB5525.1; -  
 DR InterPro: IPR002096; Cmpixl\_20kDa.  
 DR Pfam: PF01058; Oxidored\_g6; 1.  
 DR PROSITE: PS01150; COMPLEX1\_20K; 1.  
 KW NAD; Oxidoreductase; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 210 SUBUNIT NUKM OF PROTEIN NADH:UBIQUINONE  
 FT OXIDOREDUCTASE (COMPLEX I).  
 SQ SEQUENCE 210 AA; 23430 MW; C976335ECFDP686ED CRC64;  
 Query Match 57.8%; Score 37; DB 3; Length 210;  
 Best Local Similarity 63.6%; Pred. No. 23;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 FOLQRMNRKVR 12  
 Db ||||| : |||  
 Db 192 FOLQRMNRNTK 202  
 RESULT 13  
 Q45308  
 ID O45308 PRELIMINARY; PRT; 346 AA.  
 AC O45308;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE C47A10.6 protein.  
 GN C47A10.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_taxid=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Basham V.M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR EMBL: Z81484; CA503971.1; -  
 DR InterPro: IPR000344; Sra\_chemo\_Ce.  
 DR Pfam: PF02117; Sra; 1.  
 SQ SEQUENCE 346 AA; 40696 MW; 5EA98035D4E93B5B CRC64;  
 Query Match 57.8%; Score 37; DB 5; Length 346;  
 Best Local Similarity 54.5%; Pred. No. 37;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FOLQRMNRKVR 12  
 Db :||| : |||  
 Db 230 YQLEANNRNR 240  
 RESULT 14  
 Q8XJU2

ID Q8XJU2 PRELIMINARY; PRT; 565 AA.  
 AC Q8XJU2;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Arginine-trna ligase.  
 DE Arginine-trna ligase.  
 GN ARG5 OR CPE1661.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_taxid=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / TYPE A;  
 RX PubMed=11792842;  
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001 (2002).  
 DR EMBL: AP003191; BAB81367.1; -  
 DR InterPro: IPR001278; Arg\_trna-synt\_1c.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR InterPro: IPR005148; N.  
 DR InterPro: IPR001412; trna-synt\_I.  
 DR Pfam: PF03485; N-Arg; 1.  
 DR Pfam: PF00750; tRNA-synt\_id; 1.  
 DR PRINTS: PR01038; TRNASYNTHARG.  
 DR TIGRFAMs: TIGR00456; argS; 1.  
 DR PROSITE: PS00178; AA TRNA LIGASE I; 1.  
 DR PROSITE: PS00307; LECTIN LEGUME\_BETA; UNKNOWN\_1.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 565 AA; 64626 MW; 2626BD869B6ED66E CRC64;  
 Query Match 57.8%; Score 37; DB 16; Length 565;  
 Best Local Similarity 70.0%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFOLQRMNRK 10  
 Db ||||| : |||  
 Db 41 CFQAKVMNRK 50  
 RESULT 15  
 Q9SX86  
 ID Q9SX86 PRELIMINARY; PRT; 887 AA.  
 AC Q9SX86;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE F16N3.17 protein.  
 DE F16N3.17.  
 GN F16N3.17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
 RA Lee J., Liu A., Li J., Kremenetskaia I., Gonzalez A., Altafi H.,  
 RA Araujo R., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,  
 RA Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C., Rowley D.,  
 RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,  
 RA Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC F16N3 sequence.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007519; AAD46030.1; -  
 DR InterPro: IPR002048; EF-hand.  
 DR PROSITE: PS00018; EF HAND; UNKNOWN\_1.  
 SQ SEQUENCE 887 AA; 99921 MW; CEC76948485882E6 CRC64;

Tue Dec 9 06:51:40 2003

Query Match 57.8%; Score 37; DB 10; Length 887;  
Best Local Similarity 50.0%; Pred. No. 94;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLQNNMKVR 12  
| | | | | | | | | | | | | | | | | | | | | |  
Db 659 CHQIEKNERNVR 670

Search completed: February 21, 2003, 07:44:35  
Job time : 22.8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 Seconds  
(without alignments)

56.502 Million cell updates/sec

Title: US-09-743-107b-81

Perfect score: 71

Sequence: 1 CFQWQKMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 101002:\*  
1: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	100.0	12	21	AAV78081 Human lactoferrin
2	68	95.8	12	21	AAV78038 Human lactoferrin
3	68	95.8	12	21	AAV78046 Human lactoferrin
4	68	95.8	12	21	AAV78047 Human lactoferrin
5	68	95.8	13	21	AAV78037 Human lactoferrin
6	68	95.8	13	21	AAV78048 Human lactoferrin
7	68	95.8	13	21	AAV78049 Human lactoferrin
8	68	95.8	14	21	AAV78036 Human lactoferrin
9	68	95.8	14	21	AAV78050 Human lactoferrin
10	68	95.8	14	21	AAV78051 Human lactoferrin

11	68	95.8	15	17	AAV98554 Peptide for anti-u
12	68	95.8	15	21	AAV78035 Human lactoferrin
13	68	95.8	15	21	AAV78062 Human lactoferrin
14	68	95.8	15	21	AAV78063 Human lactoferrin
15	68	95.8	16	21	AAV78031 Human lactoferrin
16	68	95.8	16	21	AAV78064 Human lactoferrin
17	68	95.8	16	21	AAV78065 Human lactoferrin
18	68	95.8	17	21	AAV78034 Human lactoferrin
19	68	95.8	17	21	AAV78066 Human lactoferrin
20	68	95.8	17	21	AAV78067 Human lactoferrin
21	68	95.8	18	15	AAV69352 Human lactoferrin
22	68	95.8	18	17	AAV13397 Advanced glycosyla
23	68	95.8	18	21	AAV78033 Human lactoferrin
24	68	95.8	19	21	AAV68867 Amino acid sequenc
25	68	95.8	19	21	AAV78032 Human lactoferrin
26	68	95.8	20	13	AAV21810 Anti microbial pep
27	68	95.8	20	14	AAV44841 Lactoferrin-relate
28	68	95.8	20	15	AAV48530 Lactoferrin derive
29	68	95.8	20	15	AAV48531 Lactoferrin derive
30	68	95.8	20	15	AAV57461 Lactoferrin derive
31	68	95.8	20	15	AAV57462 Bovine lactoferrin
32	68	95.8	20	16	AAV84698 Bovine lactoferrin
33	68	95.8	20	16	AAV84699 Anti-parasitic lac
34	68	95.8	20	16	AAV80263 Anti-parasitic lac
35	68	95.8	20	16	AAV80264 Peptide for anti-u
36	68	95.8	20	17	AAV98553 Lactoferrin-derive
37	68	95.8	20	17	AAV91852 Lactoferrin-derive
38	68	95.8	20	17	AAV03045 Lactoferrin-derive
39	68	95.8	20	17	AAV90607 Lactoferrin-derive
40	68	95.8	20	17	AAV87621 Lactoferrin-derive
41	68	95.8	20	17	AAV87622 Lactoferrin-derive
42	68	95.8	20	18	AAV26150 Lactoferrin deriva
43	68	95.8	20	18	AAV14036 Anti-parasitic pep
44	68	95.8	20	18	AAV70310 Thrombus formation
45	68	95.8	20	19	AAV53224 Lactoferrin hydrol

## ALIGNMENTS

RESULT 1  
AAV78081  
ID AAV78081 standard; Peptide; 12 AA.  
XX AAV78081;  
AC AAV78081;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:81.

DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.

PD 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SS-0002562.

PR 29-DEC-1998; 98SE-0004614.

PA (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.



XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 100.0%; Score 71; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. NO. 2.3e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQKNRKVR 12  
 ID AAY78038 standard; Peptide; 12 AA.  
 AC AAY78038;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:38.  
 DE Human, lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 70; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. NO. 7.4e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQKNRKVR 12  
 ID AAY78046 standard; Peptide; 12 AA.  
 AC AAY78046;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:46.  
 DE Human, lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 7.4e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQKQNMKVR 12  
 |||||:|||||  
 Db 1 CFOWQKQNMKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;  
 XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:47.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.

XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 7.4e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQKQNMKVR 12  
 |||||:|||||  
 Db 1 CFOWQKQNMKVR 12

RESULT 5  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:37.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.

XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNRKVR 12  
 DB 2 CFQWQNNRKVR 13

## RESULT 6

AAV78048  
 ID AAV78048 standard; Peptide; 13 AA.

XX AC AAV78048;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:48.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 74; 102pp; English.

XX AA AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNRKVR 12  
 DB 2 CFQWQNNRKVR 13

## RESULT 7

AAV78049

ID AAV78049 standard; Peptide; 13 AA.

XX AC AAV78049;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:49.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 74; 102pp; English.

XX AA AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNRKVR 12  
 DB 2 CFQWQNNRKVR 13

## RESULT 8

AAV78036

ID AAV78036 standard; Peptide; 14 AA.

XX AC AAV78036;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:36.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 69; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 14 AA;  
 SQ Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQKNRKVR 12  
 DB 3 CFQWQKNRKVR 14  
 RESULT 9  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AAY78050;  
 AC AAY78050;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:50.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.

OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 75; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 14 AA;  
 SQ Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQKNRKVR 12  
 DB 3 CFQWQKNRKVR 14  
 RESULT 10  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX AAY78051;  
 AC AAY78051;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:51.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI  
 XX WPI; 2000-147388/13.  
 DR  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 XX Claim 18; Page 75; 102pp; English.  
 PS  
 XX  
 XX RAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 14 AA;  
 SQ

Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQNMNRKVR 12  
 DB 3 CFOWQNMNRKVR 14  
 |||||:|||||

RESULT 11  
 AAR98554  
 ID AAR98554 standard; Peptide; 15 AA.  
 XX  
 XX AAR98554;  
 AC  
 XX  
 XX 12-NOV-1996 (first entry)  
 DT  
 XX Peptide for anti-ulcer agent.  
 DE  
 XX anti-ulcer agent; low toxicity; stable; heat-resistant.  
 KW  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX JP08143468-A.  
 PN  
 XX  
 XX 04-JUN-1996.  
 PD  
 XX  
 XX 17-NOV-1994; 94JP-0283869.  
 PF  
 XX  
 XX 17-NOV-1994; 94JP-0283869.  
 PR  
 XX  
 XX (MORG ) MORINAGA MILK IND CO LTD.  
 PA  
 XX  
 XX WPI; 1996-318857/32.  
 DR  
 XX  
 XX Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble  
 PT  
 XX Claim 1; Page 11; 11pp; Japanese.  
 PS

XX  
 CC AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.  
 XX  
 XX Sequence 15 AA;  
 SQ

Query Match 95.8%; Score 68; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQNMNRKVR 12  
 DB 2 CFOWQNMNRKVR 13  
 |||||:|||||

RESULT 12  
 AAY78035  
 ID AAY78035 standard; Peptide; 15 AA.  
 XX  
 XX AAY78035;  
 AC  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:35.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Synthetic.  
 OS  
 XX WO200001730-A1.  
 PN  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR  
 XX 17-JUL-1998; 98SE-0002562.  
 PR  
 XX 29-DEC-1998; 98SE-0004614.  
 PR  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI  
 XX  
 XX WPI; 2000-147388/13.  
 DR  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 XX Claim 12; Page 69; 102pp; English.  
 PS  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 15 AA;  
 SQ

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
 |||||:  
 DB 4 CFQWQNNMKVR 15

## RESULT 13

AAV78062  
 ID AAY78062 standard; Peptide; 15 AA.

XX AC AAY78062;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:62.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -

XX FS Claim 15; Page 81; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
 |||||:  
 DB 4 CFQWQNNMKVR 15

## RESULT 14

AAV78063

ID AAY78063 standard; Peptide; 15 AA.

XX AC AAY78063;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:63.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -

XX FS Claim 18; Page 81; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
 |||||:  
 DB 4 CFQWQNNMKVR 15

## RESULT 15

AAV78031

ID AAY78031 standard; Peptide; 16 AA.

XX AC AAY78031;

XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:31.  
DE  
DE Human lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200001730-A1.  
XX  
XX 13-JAN-2000.  
XX  
XX 06-JUL-1999; 99WO-SE01230.  
XX  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
XX (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 11; Page 68; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
XX  
SQ Sequence 16 AA;  
Query Match 95.8%; Score 68; DB 21; Length 16;  
Best Local Similarity 91.7%; Pred. No. 9.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWQKNRKVR 12  
| | | | | | | | | | | | | |  
Db 5 CFQWQKNRKVR 16

Search completed: February 21, 2003, 07:37:14  
Job time : 28.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 Seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-81

Perfect score: 71

Sequence: 1 CFQWQNMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	18	1	US-08-204-487-3
2	68	95.8	18	2	US-08-485-948-8
3	68	95.8	18	2	US-08-628-380-8
4	68	95.8	18	2	US-08-475-055-8
5	68	95.8	20	1	US-07-755-161A-3
6	68	95.8	20	1	US-07-891-174-3
7	68	95.8	20	1	US-08-204-487-1
8	68	95.8	20	1	US-08-256-771-24
9	68	95.8	20	1	US-08-256-771-25
10	68	95.8	20	1	US-08-381-984-24
11	68	95.8	20	1	US-08-381-984-25
12	68	95.8	22	4	US-09-508-734-4
13	68	95.8	24	4	US-09-508-734-6
14	68	95.8	25	1	US-07-755-161A-10
15	68	95.8	25	1	US-07-891-174-10
16	68	95.8	25	1	US-08-204-487-7
17	68	95.8	29	4	US-09-508-734-8
18	68	95.8	36	1	US-07-755-161A-8
19	68	95.8	36	1	US-07-891-174-8
20	68	95.8	36	1	US-08-256-771-30
21	68	95.8	36	1	US-08-381-984-29
22	68	95.8	47	2	US-08-464-182A-6
23	68	95.8	47	2	US-08-406-271-6
24	68	95.8	50	2	US-08-693-274A-7
25	68	95.8	52	4	US-09-017-043A-3
26	68	95.8	53	2	US-08-464-182A-5
27	68	95.8	53	2	US-08-406-271-5

28 68 95.8 54 2 US-08-464-182A-2 Sequence 2, Appli  
29 68 95.8 54 2 US-08-406-271-2 Sequence 2, Appli  
30 68 95.8 694 3 US-08-724-586-2 Sequence 2, Appli  
31 68 95.8 694 4 US-09-421-632-2 Sequence 2, Appli  
32 68 95.8 694 4 US-09-932-190-2 Sequence 2, Appli  
33 68 95.8 705 2 US-08-655-640-2 Sequence 2, Appli  
34 68 95.8 708 2 US-08-655-640-4 Sequence 4, Appli  
35 68 95.8 711 1 US-08-154-019-4 Sequence 4, Appli  
36 68 95.8 711 1 US-08-461-333-4 Sequence 4, Appli  
37 68 95.8 711 3 US-08-464-167-4 Sequence 4, Appli  
38 68 95.8 711 3 US-09-158-313-4 Sequence 4, Appli  
39 68 95.8 711 4 US-08-476-798-4 Sequence 4, Appli  
40 65 91.5 711 1 US-08-145-681-2 Sequence 2, Appli  
41 65 91.5 711 1 US-08-250-308-2 Sequence 2, Appli  
42 65 91.5 711 1 US-08-453-703-2 Sequence 2, Appli  
43 65 91.5 711 2 US-08-456-106-2 Sequence 2, Appli  
44 65 91.5 711 3 US-08-456-108-2 Sequence 2, Appli  
45 65 91.5 711 4 US-09-265-577-2 Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"



OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 95.8%; Score 68; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQKNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQKNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQKNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

1 CORRESPONDENCE ADDRESS:  
2 ADDRESSEE: Klauber & Jackson  
3 STREET: 411 Hackensack Avenue  
4 CITY: Hackensack  
5 STATE: New Jersey  
6 COUNTRY: USA  
7 ZIP: 07601  
8  
9 COMPUTER READABLE FORM:  
10 MEDIUM TYPE: Floppy disk  
11 COMPUTER: IBM PC compatible  
12 OPERATING SYSTEM: PC-DOS/MS-DOS  
13 SOFTWARE: Patent in Release #1.0, Version #1.25  
14 CURRENT APPLICATION DATA:  
15 APPLICATION NUMBER: US/08/475,055  
16 FILING DATE:  
17 CLASSIFICATION:  
18 PRIOR APPLICATION DATA:  
19 APPLICATION NUMBER: 08/485,948  
20 FILING DATE:  
21 APPLICATION NUMBER: 08/488,217  
22 FILING DATE: JUNE 7, 1995  
23 APPLICATION NUMBER: 08/418,642  
24 FILING DATE: APRIL 7, 1995  
25 CLASSIFICATION:  
26 ATTORNEY/AGENT INFORMATION:  
27 NAME: Jackson Esq., David A.  
28 REGISTRATION NUMBER: 26,742  
29 REFERENCE/DOCKET NUMBER: 947-1-008A  
30 TELECOMMUNICATION INFORMATION:  
31 TELEPHONE: 201 487-5800  
32 TELEFAX: 201 343-1684  
33  
34 INFORMATION FOR SEQ ID NO: 8:  
35 SEQUENCE CHARACTERISTICS:  
36 LENGTH: 18 amino acids  
37 TYPE: amino acid  
38 STRANDEDNESS: single  
39 TOPOLOGY: linear  
40 MOLECULE TYPE: peptide  
41 DESCRIPTION: LF-C1, 8-25  
42 HYPOTHETICAL: NO  
43 ANTI-SENSE: NO  
44 FRAGMENT TYPE: internal  
45 US-08-475-055-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNMKVR 12  
| | | | |  
Db 1 CFQWQKNMKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

1 OPERATING SYSTEM: MS-DOS  
2 SOFTWARE: DisplayWrite  
3 CURRENT APPLICATION DATA:  
4 APPLICATION NUMBER: US/07/755,161A  
5 FILING DATE: 19910905  
6 CLASSIFICATION: 530  
7 PRIOR APPLICATION DATA:  
8 APPLICATION NUMBER:  
9 FILING DATE:  
10 ATTORNEY/AGENT INFORMATION:  
11 NAME: Warren M. Cheek Jr.  
12 REGISTRATION NUMBER: 33,367  
13 REFERENCE/DOCKET NUMBER:  
14 TELECOMMUNICATION INFORMATION:  
15 TELEPHONE: 202-371-8850  
16 TELEFAX: 202-371-8856  
17  
18 INFORMATION FOR SEQ ID NO: 3:  
19 SEQUENCE CHARACTERISTICS:  
20 LENGTH: 20 amino acids  
21 TYPE: AMINO ACID  
22 STRANDEDNESS: single  
23 TOPOLOGY: linear  
24 MOLECULE TYPE:  
25 HYPOTHETICAL:  
26 ANTI-SENSE:  
27 FRAGMENT TYPE:  
28 ORIGINAL SOURCE:  
29 ORGANISM:  
30 STRAIN:  
31 INDIVIDUAL ISOLATE:  
32 DEVELOPMENTAL STAGE:  
33 HAPLOTYPE:  
34 TISSUE TYPE:  
35 CELL TYPE:  
36 CELL LINE:  
37 ORGANELLE:  
38 IMMEDIATE SOURCE:  
39 LIBRARY:  
40 CLONE:  
41 POSITION IN GENOME:  
42 CHROMOSOME/SEGMENT:  
43 MAP POSITION:  
44 UNITS:  
45 FEATURE:  
46 NAME/KEY: modified site  
47 LOCATION: 2  
48 IDENTIFICATION METHOD:  
49 OTHER INFORMATION: /note= "thiol group of  
50 OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
51 OTHER INFORMATION: thiol group of Cys residue at location 19"  
52 FEATURE:  
53 NAME/KEY: modified site  
54 LOCATION: 19  
55 IDENTIFICATION METHOD:  
56 OTHER INFORMATION: /note= "thiol group of  
57 OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
58 OTHER INFORMATION: thiol group of Cys residue at location 2"  
59 PUBLICATION INFORMATION:  
60 AUTHORS:  
61 TITLE:  
62 JOURNAL:  
63 VOLUME:  
64 ISSUE:  
65 PAGES:  
66 DATE:  
67 DOCUMENT NUMBER:  
68 FILING DATE:  
69 PUBLICATION DATE:  
70 RELEVANT RESIDUES IN SEQ ID NO:  
71 US-07-755-161A-3

Query Match 95.8%; Score 68; DB 1; Length 20;

Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12  
Db 2 CFQWQKNRKVR 13

## RESULT 6

US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Fonack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/891,174

FILING DATE: 29-MAY-1992

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/755,161

FILING DATE: 05-SEP-1991

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM:

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-07-891-174-3  
Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12

Db 2 CFQWQKNRKVR 13

## RESULT 7

US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGRAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,487  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A.  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
|||:|||||  
Db 2 CFQWRNRKVR 13

RESULT 8  
US-08-256-771-24  
; Sequence 24, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,771  
; FILING DATE: July 22, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "Cys residues are linked by  
; OTHER INFORMATION: disulfide bond"

US-08-256-771-24  
Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
|||:|||||  
Db 2 CFQWRNRKVR 13

RESULT 9  
US-08-256-771-25  
; Sequence 25, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,771  
; FILING DATE: July 22, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "Cys residues are protected to  
; OTHER INFORMATION: prevent disulfide bond"  
US-08-256-771-25

Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
|||:|||||  
Db 2 CFQWRNRKVR 13

RESULT 10  
US-08-381-984-24



; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-508-734-4

Query Match 95.8%; Score 68; DB 4; Length 22;  
Best Local Similarity 91.7%; Pred. No. 5.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQKNRKVR 12  
Db 2 CFQWQKNRKVR 13

## RESULT 13

US-09-508-734-6  
; Sequence 6, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508.734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KE99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: KopatentIn 1.71  
; SEQ ID NO 6  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-508-734-6

Query Match 95.8%; Score 68; DB 4; Length 24;  
Best Local Similarity 91.7%; Pred. No. 6e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQKNRKVR 12  
Db 3 CFQWQKNRKVR 14

## RESULT 14

US-07-755-161A-10  
; Sequence 10, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/755.161A  
; FILING DATE: 19910905  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; Query Match 95.8%; Score 68; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 6.3e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CFQWQKNRKVR 12  
Db 4 CFQWQKNRKVR 15  
US-07-755-161A-10  
; RELEVANT RESIDUES IN SEQ ID NO:  
; Query Match 95.8%; Score 68; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 6.3e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CFQWQKNRKVR 12  
Db 4 CFQWQKNRKVR 15

RESULT 15  
US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10  
; Query Match 95.8%; Score 68; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 6.3e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
; Qy 1 CFQWQKNNRKVR 12  
; Db 4 CFQWQKNNRKVR 15  
; Search completed: February 21, 2003, 07:50:36  
; Job time : 9.7 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-81  
Perfect score: 71  
Sequence: 1 CFQWQNNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31059816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	ID	Description
1	68	95.8	15	US-09-798-869-2
2	68	95.8	25	US-09-798-869-20
3	68	95.8	694	US-10-023-096-2
4	63	88.7	15	US-09-798-869-6
5	51	71.8	15	US-09-798-869-3
6	51	71.8	25	US-09-798-869-23
7	46	64.8	15	US-09-798-869-7
8	45	63.4	15	US-09-798-869-4
9	45	63.4	25	US-09-798-869-22
10	42	59.2	15	US-09-798-869-8
11	42	59.2	15	US-09-798-869-29
12	42	59.2	15	US-09-798-869-30
13	37	52.1	40	US-09-864-761-46393
14	37	52.1	747	US-10-066-500-58
15	37	52.1	747	US-10-002-796-58
16	37	52.1	747	US-10-066-273-58
17	37	52.1	747	US-10-066-494-58
18	36	50.7	77	US-09-864-761-41002
19	36	50.7	239	US-09-864-761-37353

20	35	49.3	15	9	US-09-798-869-5	Sequence 5, Appli
21	35	49.3	319	10	US-09-815-242-10267	Sequence 10267, A
22	35	49.3	434	10	US-09-815-242-10441	Sequence 10441, A
23	35	49.3	489	9	US-09-888-320-2	Sequence 2, Appli
24	34	47.9	21	10	US-09-864-761-47985	Sequence 47985, A
25	34	47.9	56	10	US-09-864-761-44710	Sequence 44710, A
26	34	47.9	62	10	US-09-815-242-12129	Sequence 12129, A
27	34	47.9	62	10	US-09-815-242-13026	Sequence 13026, A
28	34	47.9	351	10	US-09-853-625B-16	Sequence 16, Appli
29	34	47.9	447	9	US-10-081-309-2	Sequence 2, Appli
30	34	47.9	723	9	US-10-133-912-18	Sequence 18, Appli
31	34	47.9	723	9	US-10-133-912-19	Sequence 19, Appli
32	34	47.9	760	10	US-09-265-606-2	Sequence 2, Appli
33	34	47.9	933	9	US-10-001-189-60	Sequence 60, Appli
34	34	47.9	1013	9	US-10-028-072-38	Sequence 38, Appli
35	34	47.9	1013	9	US-10-121-049-38	Sequence 38, Appli
36	34	47.9	1013	9	US-10-123-904-38	Sequence 38, Appli
37	34	47.9	1013	9	US-10-140-470-38	Sequence 38, Appli
38	34	47.9	1013	9	US-10-175-746-38	Sequence 38, Appli
39	34	47.9	1013	9	US-10-176-918-38	Sequence 38, Appli
40	34	47.9	1013	9	US-10-176-921-38	Sequence 38, Appli
41	34	47.9	1013	9	US-10-137-865-38	Sequence 38, Appli
42	34	47.9	1013	9	US-10-140-474-38	Sequence 38, Appli
43	34	47.9	2273	10	US-09-995-542-12	Sequence 12, Appli
44	34	47.9	2310	10	US-09-995-542-10	Sequence 10, Appli
45	34	47.9	2332	9	US-09-957-641-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2  
Query Match 95.8%; Score 68; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 2.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWQNNRKVR 12  
|||:|||||  
Db 3 CFQWQNNRKVR 14  
|||:|||||  
RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON



APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 95.8%; Score 68; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 4.6e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNNRKVR 12  
| | | | |  
Db 3 CFQWQNNRKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 95.8%; Score 68; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNNRKVR 12  
| | | | |  
Db 22 CFQWQNNRKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 88.7%; Score 63; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.00018;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQNNRKVR 12  
| | | | |  
Db 3 CFQWQNNRKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 71.8%; Score 51; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.016;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWQKNMRKV 11  
|:|:|:|:|:|:  
Db 3 CYOWQRRMRKL 13

RESULT 6  
US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 71.8%; Score 51; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.027;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWQKNMRKV 11  
|:|:|:|:|:|:  
Db 3 CYOWQRRMRKL 13

RESULT 7  
US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 64.8%; Score 46; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.11;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQKNMRKV 11  
|:|:|:|:|:|:  
Db 3 CYOWQRRMRKL 13

RESULT 8  
US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 63.4%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.16;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQKNMRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQNRKV 13

RESULT 9  
US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 63.4%; Score 45; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.25;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQKNMRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQNRKV 13

## RESULT 10

US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.48;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKNRKV 11  
| : || || ||  
DB 3 CLRQWQWNRKV 13

## RESULT 11

US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.48;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQKNRKV 11  
| : || || ||  
DB 3 CFRQWQWNRKV 13

## RESULT 12

US-09-798-869-30

; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.48;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQKNRKV 11  
| : || || ||  
DB 3 CFRQWQWNRKV 13

## RESULT 13

US-09-864-761-46393  
; Sequence 46393, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46393
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018719.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: N80993.1, EVALUATE 9.00e-11
US-09-864-761-46393

```

```

Query Match. 52.1%; Score 37; DB 10; Length 40;
Best Local Similarity 75.0%; Pred. No. 8;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
QY 2 FOWQKNWR 9
```

```
Db 21 FOWDKNWR 28
```

# RESULT 14

```
US-10-066-500-58
; Sequence 58, Application US/10066500
; Patent No. US20020177165A1

```

## GENERAL INFORMATION:

```

; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Botstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Eaton
; APPLICANT: Napoleon Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tunas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3130RIC7
; CURRENT APPLICATION NUMBER: US/10/066,500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115

```

```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-03-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106032
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/109304
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/139695
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/145070
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/149396
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 08/918874
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 08/933821
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 08/960507
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 09/114844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: 09/136801
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136804
; PRIOR FILING DATE: 1998-08-19

```

PRIOR APPLICATION NUMBER: 09/136828  
PRIOR FILING DATE: 1998-08-19  
PRIOR APPLICATION NUMBER: 09/158342  
PRIOR FILING DATE: 1998-09-21  
PRIOR APPLICATION NUMBER: 09/180997  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 09/202088  
PRIOR FILING DATE: 1998-12-08  
PRIOR APPLICATION NUMBER: 09/254311  
PRIOR FILING DATE: 1999-03-03  
PRIOR APPLICATION NUMBER: 09/254460  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/254465  
PRIOR FILING DATE: 1999-03-05  
PRIOR APPLICATION NUMBER: 09/284663  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 09/332928  
PRIOR FILING DATE: 1999-06-14  
PRIOR APPLICATION NUMBER: 09/332929  
PRIOR FILING DATE: 1999-06-14  
PRIOR APPLICATION NUMBER: 09/333075  
PRIOR FILING DATE: 1999-06-14  
PRIOR APPLICATION NUMBER: 09/333077  
PRIOR FILING DATE: 1999-06-14  
PRIOR APPLICATION NUMBER: 09/380137  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/380138  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/380139  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/403296  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: 09/403297  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: 09/423741  
PRIOR FILING DATE: 1999-11-10  
PRIOR APPLICATION NUMBER: 09/423844  
PRIOR FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: 09/522342  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: 09/548815  
PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: 09/664610  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/665350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/709238  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 09/767609  
PRIOR FILING DATE: 2001-01-22  
PRIOR APPLICATION NUMBER: 09/802706  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 09/808689  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 09/866028  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: 09/870574  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 09/872035  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 09/886342  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: PCT/US98/14552  
PRIOR FILING DATE: 1998-07-14  
PRIOR APPLICATION NUMBER: PCT/US98/18824  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: PCT/US98/19093  
PRIOR FILING DATE: 1998-09-14  
PRIOR APPLICATION NUMBER: PCT/US98/19330  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: PCT/US98/19437  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: PCT/US98/24855

PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: PCT/US98/25108  
PRIOR FILING DATE: 1998-12-01  
PRIOR APPLICATION NUMBER: PCT/US98/25190  
PRIOR FILING DATE: 1998-11-25  
PRIOR APPLICATION NUMBER: PCT/US99/05028  
PRIOR FILING DATE: 1999-03-08  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/20111  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 52.1%; Score 37; DB 9; length 747;  
Best Local Similarity 45.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQKMKRV 11  
| : | | | : :  
Db 311 CWRQINSRI 321

## RESULT 15

US-10-002-796-58  
; Sequence 58, Application US/10002796  
; Publication No. US20030032057A1

## GENERAL INFORMATION:

; APPLICANT: Avi J. Ashkenazi  
; APPLICANT: Kevin P. Baker  
; APPLICANT: David A. Botstein  
; APPLICANT: Luc Desnoyers  
; APPLICANT: Dan L. Eaton  
; APPLICANT: Napoleone Ferrara  
; APPLICANT: Sherman Fong  
; APPLICANT: Wei-Qiang Gao  
; APPLICANT: Hanspeter Gerber  
; APPLICANT: Mary E. Gerritsen  
; APPLICANT: Audrey Goddard  
; APPLICANT: Paul J. Godowski  
; APPLICANT: Austin L. Gurney  
; APPLICANT: Ivar J. Kljavin  
; APPLICANT: Jennie P. Mather  
; APPLICANT: Mary A. Napier  
; APPLICANT: James Pan  
; APPLICANT: Nicholas F. Paoni  
; APPLICANT: Margaret Ann Roy  
; APPLICANT: Timothy A. Stewart  
; APPLICANT: Daniel Tumas  
; APPLICANT: Colin K. Watanabe  
; APPLICANT: P. Mickey Williams  
; APPLICANT: William I. Wood  
; APPLICANT: Zemin Zang  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3130R1C1  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062285  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062816  
; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063082  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/063329  
; PRIOR FILING DATE: 1997-10-27  
; PRIOR APPLICATION NUMBER: 60/063733  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066840  
; PRIOR FILING DATE: 1997-11-25  
; PRIOR APPLICATION NUMBER: 60/069694  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 60/074086  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/074092  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/095998  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/097000  
; PRIOR FILING DATE: 1998-08-18  
; PRIOR APPLICATION NUMBER: 60/099601  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099811  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100858  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101922  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/106032  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/109304  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/125778  
; PRIOR FILING DATE: 1998-03-23  
; PRIOR APPLICATION NUMBER: 60/139695  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/145070  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/149396  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 08/918874  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 08/933821  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 08/960507  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 09/114844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: 09/136801  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136804  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136828  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/158342  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: 09/180997  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 09/202088  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: 09/254311

; PRIOR FILING DATE: 1999-03-03  
; PRIOR APPLICATION NUMBER: 09/254460  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: 09/254465  
; PRIOR FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: 09/284663  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 09/332928  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/332929  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333075  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333077  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380139  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/403296  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/403297  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/423741  
; PRIOR FILING DATE: 1999-11-10  
; PRIOR APPLICATION NUMBER: 09/423844  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 09/522342  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 09/548815  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: 09/664610  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/665350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/709238  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/767609  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 09/802706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 09/808689  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/866028  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 09/870574  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 09/872035  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/886342  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: PCT/US98/14552  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US98/18824  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/US98/19093  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: PCT/US98/19330  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: PCT/US98/19437  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/24855  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: PCT/US98/25108  
; PRIOR FILING DATE: 1998-12-01  
; PRIOR APPLICATION NUMBER: PCT/US98/25190  
; PRIOR FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02

; PRIOR APPLICATION NUMBER: PCT/US99/20111  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/28301

Query Match 52.1%; Score 37; DB 9; Length 747;  
Best Local Similarity 45.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWQXMRKV 11  
| : | | | : :  
Db 311 CVRWQINSRI 321

Search completed: February 21, 2003, 08:08:07  
Job time : 10.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-81

Perfect score: 71

Sequence: 1 CFQWQKNMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	711	1	TFHUL
2	54	76.1	33	2	lactotransferrin p
3	51	71.8	708	2	lactoferrin - shee
4	45	63.4	707	1	lactoferrin - goat
5	45	63.4	4568	2	lactoferrin precu
6	43	60.6	681	2	dynain beta heavy
7	42	59.2	511	2	hypothetical prote
8	42	59.2	584	2	hypothetical prote
9	41	57.7	275	2	hypothetical prote
10	39	54.9	283	2	sugar ABC transpor
11	39	54.9	298	2	hypothetical prote
12	39	54.9	1135	2	AD2346
13	39	54.9	1274	2	phytichrome C - so
14	38	53.5	206	2	G74018
15	38	53.5	206	2	H97451
16	38	53.5	206	2	AB2670
17	38	53.5	208	2	AG3441
18	38	53.5	211	2	D82109
19	38	53.5	289	2	F90580
20	38	53.5	282	2	G86403
21	38	53.5	317	2	T47233
22	38	53.5	393	2	D71876
23	38	53.5	428	2	F81660
24	38	53.5	759	2	G86506
25	38	53.5	759	2	T28820
26	38	53.5	1432	2	B85431
27	37	52.1	99	2	C97838
28	37	52.1	316	2	D71535
29	37	52.1	334	1	KHRTL

30	37	52.1	334	1	KHMSL
31	37	52.1	361	2	T29571
32	37	52.1	367	2	G97649
33	37	52.1	393	2	E64639
34	37	52.1	496	2	T50146
35	37	52.1	570	2	T46261
36	37	52.1	583	2	T03470
37	37	52.1	749	2	A45687
38	36	50.7	120	2	T33605
39	36	50.7	306	1	A39654
40	36	50.7	338	2	T04734
41	36	50.7	415	2	C71467
42	36	50.7	431	2	S50977
43	36	50.7	476	2	B9547
44	36	50.7	500	2	T49978
45	36	50.7	518	2	B84514

#### ALIGNMENTS

##### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R:Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R:Rev. M.M.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148 'T', 150-422 'C', 424-711 <REV>

A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Feng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A:Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A:Experimental source: placenta

A>Note: sequence extracted from NCBI backbone (NCBIP:122202)

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A:Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <STI>

A:Accession: S20841

A:Molecule type: Protein

A:Residues: 20-28, 'X', 30-31 <ST2>



R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A:Reference number: S07160; PMID:88001031; PMID:3477300  
 A:Accession: S07160  
 A:Molecule type: mRNA  
 A:Residues: 436-487, 'A', 489-711 <RAD>  
 A:Cross-references: EMBL:M18642; NID:G186815; PID:AAA66665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A:Reference number: A61169; PMID:91235214; PMID:1674448  
 A:Accession: A61169  
 A:Molecule type: mRNA  
 A>Status: not compared with conceptual translation  
 A:Residues: 3-701, 'SWPKPVN' <PAN>  
 A:Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1984  
 A>Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A:Reference number: A31000; PMID:85076667; PMID:6510420  
 A:Accession: A31000  
 A:Molecule type: protein  
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-400  
 A:Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Bakkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A:Reference number: S74119; PMID:97054624; PMID:8898921  
 A:Accession: S74119  
 A:Molecule type: protein  
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A:Experimental source: neutrophil granulocytes  
 C:Genetics:  
 A:Gene: GDB:ITF  
 A:Cross-references: GDB:119368; OMIM:150210  
 A:Map position: 3q21-3q23  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-71/Product: lactotransferrin repeat homology <MAT>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status experimental  
 F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #status experimental

Query Match 95.8%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0054;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 12  
 Db 39 CFQWQKNRKVR 50

RESULT 2  
 S52107  
 lactoferrin - sheep (fragment)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C:Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet aggregation  
 A:Reference number: S52107; PMID:95127729; PMID:7827104  
 A:Accession: S52107  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-33 <QIA>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication

Query Match 76.1%; Score 54; DB 2; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.0069;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 11  
 Db 19 CYQWQKNRKRL 29

RESULT 3  
 JC2323  
 lactoferrin - goat  
 C:Species: Capra aegagrus hircus (domestic goat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C:Accession: JC2323  
 R;Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A:Reference number: JC2323; MUID:94380047; PMID:8093048  
 A:Accession: JC2323  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <LEP>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:359-696/Domain: transferrin repeat homology <TRH2>  
 F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.8%; Score 51; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.52;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 11  
 Db 38 CYQWQKNRKRL 48

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N:Alternate names: lactotransferrin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A>Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory gland  
 A:Reference number: A92596; MUID:87280033; PMID:3611056  
 A:Accession: A28438  
 A:Molecule type: mRNA  
 A:Residues: 3-707 <PEN>  
 A:Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A>Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A:Reference number: A41205; MUID:92042099; PMID:1939212  
 A:Accession: A41205  
 A:Molecule type: DNA  
 A:Residues: 1-15 <LIU>  
 A:Cross-references: GB:M74778  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-707/Product: lactotransferrin repeat homology <MAT>  
 F:358-695/Domain: transferrin repeat homology <TRH2>  
 F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.4%; Score 45; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 5.9;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 11  
 Db 37 CLRQWQKNRKVR 47

RESULT 5  
 T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 C:Accession: T08030  
 R: Mitchell, D.R.; Brown, K.S.  
 J. Cell Sci. 107, 635-644, 1994  
 A:Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
 A:Sequence number: 216302; MUID:94274778; PMID:8006077  
 A:Accession: T08030  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-4568 <MIT>  
 A:Cross-references: EMBL:U02963; NID:g409965; PIDN:AAA19956.1; PID:g514215  
 A:Experimental source: strain 21gr  
 C:Genetics:  
 A:Gene: ODA4  
 A:Map position: IX  
 A:Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;  
 3334/3; 3686/3; 3882/3; 4240/3  
 C:Superfamily: dynein heavy chain, ciliary  
 C:Keywords: nucleotide binding; P-loop  
 F:1919-1926/Region: nucleotide-binding motif A (P-loop)  
 F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
 F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 63.4%; Score 45; DB 2; Length 4568;  
 Best Local Similarity 50.0%; Pred. No. 39;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
 ||||| : : :  
 Db 1852 CFQWQSLRVIQ 1863

RESULT 6  
 T19429  
 hypothetical protein C24H11.8 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T19429  
 R: Lloyd, C.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19123  
 A:Accession: T19429  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-681 <WIL>  
 A:Cross-references: EMBL:Z81475; PIDN:CAB03914.1; GSPDB:GN00021; CESP:C24H11.8  
 A:Experimental source: clone C24H11  
 C:Genetics:  
 A:Gene: CESP:C24H11.8  
 A:Map position: 3  
 A:Introns: 18/3; 65/3; 108/1; 138/1; 177/3; 256/2; 303/2; 343/1; 409/3; 484/3; 532/2; 56

Query Match 60.6%; Score 43; DB 2; Length 681;  
 Best Local Similarity 70.0%; Pred. No. 13;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQKMRKVR 12  
 : ||||| : ||  
 Db 192 RWQKRRRVR 201

RESULT 7  
 A08058  
 hypothetical protein STY3070 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhimurium  
 C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001

C:Accession: A08058  
 R: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: A08058  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-511 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:g16504016; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY3070

Query Match 59.2%; Score 42; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
 ||||| : : :  
 Db 350 CFQWQKMRKVR 361

RESULT 8  
 C84325  
 hypothetical protein Vng1732c [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: C84325  
 R: Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabłoński, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: C84325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-584 <STO>  
 A:Cross-references: GB:AE004437; NID:g10581192; PIDN:AAG19967.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG1732C

Query Match 59.2%; Score 42; DB 2; Length 584;  
 Best Local Similarity 50.0%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKMRKVR 12  
 ||||| : : :  
 Db 445 CFTWRKMRKVR 456

RESULT 9  
 T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T22597  
 R: Dobson, R.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19587  
 A:Accession: T22597  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-275 <WIL>  
 A:Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4  
 A:Experimental source: clone F53H4  
 C:Genetics:  
 A:Gene: CESP:F53H4.4  
 A:Map position: X

```

A;Introns: 67/1; 153/1
C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match          57.7%; Score 41; DB 2; Length 275;
Best Local Similarity 63.6%; Pred. No. 1;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQKNRKVR 12
    |||: |||:
Db 262 FQWKSMRKTR 272

RESULT 10
D72378
sugar ABC transporter, permease protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D72378
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72378
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <ARN>
A;Cross-references: GB:AE001721; GB:AE000512; NID:G4980922; PIDN:AAD35505.1; PID:G498092
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0420
C;Superfamily: maltose transport protein malG

Query Match          54.9%; Score 39; DB 2; Length 283;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WQKNRKVR 12
    |||: |||:
Db 7 WQKNSEKIR 15

RESULT 11
AD2346
hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AD2346
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yanada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2346
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA076022.1; PID:gl7133459; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr4323

Query Match          54.9%; Score 39; DB 2; Length 298;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQKNRKVR 10
    |||: |||:
Db 163 FHWQNRKIR 171

RESULT 12
TI4803
phytochrome C - sorghum
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: TI4803
R;Childs, K.L.; Miller, F.R.; Cordonnier-Pratt, M.M.; Pratt, L.H.; Morgan, P.W.; Mullet
submitted to the EMBL Data Library, April 1996
A;Reference number: Z18186
A;Accession: TI4803
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1135 <CHI>
A;Cross-references: EMBL:U56731; NID:gl800218; PID:gl800219
C;Genetics:
A;Gene: PHYC
C;Superfamily: phytochrome; phytochrome homology
C;Keywords: chromoprotein; dimer; photoreceptor; phytochromobilin; transcription regula
F:65-581/Domain: phytochrome homology <PHY>
F:32/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match          54.9%; Score 39; DB 2; Length 1135;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKNRKVR 11
    |||: |||:
Db 775 CLEWKNKMQKI 785

RESULT 13
T04018
hypothetical protein F17A8.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C;Accession: T04018
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.;
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15184
A;Accession: T04018
A;Molecule type: DNA
A;Residues: 1-1274 <BEV>
A;Cross-references: EMBL:AL049482
A;Experimental source: cultivar Columbia; BAC clone F17A8
C;Genetics:
A;Map position: 4
A;Introns: 55/3; 801/1
A;Note: F17A8.60

Query Match          54.9%; Score 39; DB 2; Length 1274;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQKNRKVR 11
    |||: |||:
Db 968 FHWQKNIRKI 977

RESULT 14
H97451
pyridoxamine 5'-phosphate oxidase (AFL79611) [imported] - Agrobacterium tumefaciens (st
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C;Accession: H97451
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; PMID:11743194
A;Accession: H97451
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <KUR>

```

A;Cross-references: GB:AE007869; PIDN:AAK86569.1; PID:g15155733; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C\_1381  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 53.5%; Score 38; DB 2; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWQNNMKVR 12  
 |||:|:  
 Db 88 CFHWKSLRRQVR 99

## RESULT 15

AB2670  
 pyridoxamine 5'-phosphate oxidase [imported] - Agrobacterium tumefaciens (strain C58, Du  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
 C;Accession: AB2670  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; PMID:11743193  
 A;Accession: AB2670  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL41776.1; PID:g17739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: pdxH  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 53.5%; Score 38; DB 2; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWQNNMKVR 12  
 |||:|:  
 Db 88 CFHWKSLRRQVR 99

Search completed: February 21, 2003, 07:47:53  
 Job time : 9.65 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 Seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-81  
Perfect score: 71  
Sequence: 1 CFQWQKMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	711	1	TRFL_HUMAN
2	51	71.8	708	1	TRFL_CAMDR
3	51	71.8	708	1	TRFL_CAPHI
4	45	63.4	707	1	TRFL_MOUSE
5	45	63.4	4568	1	DYHB_CHLRE
6	43	60.6	695	1	TRFL_HORSE
7	39	54.9	1135	1	PHYC_SORBI
8	38	53.5	62	1	RL28_THETN
9	38	53.5	211	1	LOUB_VIBCH
10	38	53.5	428	1	SYH_CHLMU
11	37	52.1	316	1	NORC_CHLTR
12	37	52.1	334	1	CATL_MOUSE
13	37	52.1	334	1	CATL_RAT
14	37	52.1	455	1	YKYL_CAEEL
15	37	52.1	496	1	MS81_SCHPO
16	37	52.1	749	1	VP4_ROTGA
17	36	50.7	292	1	NLA_DROME
18	36	50.7	329	1	BUB2_YEAST
19	36	50.7	329	1	CATK_RAT
20	36	50.7	962	1	YBX7_SCHPO
21	36	50.7	4568	1	DYHC_CAEEL
22	35	49.3	146	1	RPOB_LIBAF
23	35	49.3	211	1	FDOI_ECOLI
24	35	49.3	238	1	PELX_ERWCA
25	35	49.3	238	1	YBM9_SCHPO
26	35	49.3	267	1	Y125_MYCCA
27	35	49.3	319	1	R1N4_ECOLI
28	35	49.3	374	1	PEL2_ERWCA
29	35	49.3	434	1	ACEA_ECOLI
30	35	49.3	434	1	ACEA_SALTY
31	35	49.3	475	1	YBEA_HAEIN
32	35	49.3	485	1	GLCA_BACST
33	35	49.3	699	1	EST1_YEAST
					P02788 homo sapien
					Q9tum0 camelus dro
					Q29477 capra hircu
					P08071 mus musculu
					Q39565 chlamydomon
					O77811 equus cabal
					P93528 sorghum bic
					Q8r9u1 thermoaer
					P57070 vibrio chol
					Q9pij9 chlamydia m
					O84281 chlamydia t
					P06797 mus musculu
					P07154 rattus norv
					Q19910 caenorhabdi
					Q9ute7 schizosacch
					Q04916 rotavirus (
					Q9x2l8 drosophila
					P26448 saccharomyc
					Q35186 rattus norv
					Q10201 schizosacch
					Q19020 caenorhabdi
					P41187 liberibacte
					P32174 escherichia
					P16530 erwinia car
					Q10333 schizosacch
					P53661 mycoplasma
					P37146 escherichia
					P11431 erwinia car
					P05313 escherichia
					P51066 salmonella
					P44693 haemophilus
					O08328 bacillus st
					P17214 saccharomyc

34	35	49.3	704	1	TRFL_PIG	P14632 sus scrofa
35	35	49.3	708	1	TRFL_BUBBU	O77698 bubalus bub
36	35	49.3	966	1	VIA_BBMV	Q00020 broad bean
37	35	49.3	1095	1	PIPA_DROME	P13217 drosophila
38	35	49.3	1137	1	PHYC_ORISA	Q9zw19 oryza sativ
39	35	49.3	2096	1	BP28_DROME	Q9vm75 drosophila
40	34	47.9	60	1	RL28_BACST	Q23374 bacillus st
41	34	47.9	62	1	RL28_LISMO	Q23aj2 listeria mo
42	34	47.9	62	1	RL28_STAAM	Q99up4 staphylococ
43	34	47.9	211	1	YRBC_ECOLI	P45390 escherichia
44	34	47.9	270	1	PDXH_WYXXA	P21159 myxococcus
45	34	47.9	275	1	IL2A_BOVIN	P12342 bos taurus

## ALIGNMENTS

RESULT 1  
ID TRFL\_HUMAN STANDARD; PRT; 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q16789; O00756; Q9H1Z3; Q95KZ4;  
AC Q96KZ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactoferrin precursor (lactoferrin) [Contains: Lactoferrin A;  
DE Lactoferrin B; Lactoferrin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cho Y.Y.;  
RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Conneely O.M.;  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary Gland;  
RA Liang Q., Jimenez-Flores R., Richardson T.;  
RL "Molecular cloning and sequence analysis of human lactoferrin.";  
RN Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RL "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences";  
RN Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RL "cDNA cloning and sequence analysis of human lactoferrin";  
RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE-Mammary gland.  
 RX MEDLINE=90326649; PubMed=2374734;  
 RA Powell M.J., Ogden J.E.;  
 RT "Nucleotide sequence of human lactoferrin cDNA.";  
 RL Nucleic Acids Res. 18:4013-4013 (1990).  
 RN [19]  
 RP SEQUENCE OF 20-711.  
 RX MEDLINE=85076667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legrand D., Spik G., Montreuil J., Jolles P.;  
 RT "Human lactotransferrin: amino acid sequence and structural  
 comparisons with other transferrins.";  
 RL Eur. J. Biochem. 145:659-666 (1984).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RX MEDLINE=82046817; PubMed=6794640;  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactotransferrin sequence. Study and  
 alignment of the cyanogen bromide fragments and characterization of  
 N- and C-terminal domains.";  
 RL Biochim. Biophys. Acta 670:243-254 (1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RX MEDLINE=82262043; PubMed=7049727;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 lactotransferrin.";  
 RL FEBS Lett. 142:107-110 (1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RX MEDLINE=88001031; PubMed=3477300;  
 RA Rado T.A., Wei X., Benz E.J. Jr.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 expression of mRNA during normal and leukemic myelopoiesis.";  
 RL Blood 70:989-993 (1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Farnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
 RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
 RA Dragan Y., Glacaloni J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
 RA Sagripanti J.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=90064528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 and refinement at 2.8-A resolution.";  
 RL J. Mol. Biol. 209:711-734 (1989).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA Haridas M., Anderson B.F., Baker E.N.;  
 RT "Structure of human diferric lactoferrin refined at 2.2-A  
 resolution.";  
 RL Acta Crystallogr. D 51:629-646 (1995).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
 RX MEDLINE=97156796; PubMed=9003186;  
 RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
 RA Baker E.N.;  
 RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
 binding properties and crystal structure of the histidine-  
 253->-methionine mutant.";  
 RL Biochemistry 36:341-346 (1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=99190892; PubMed=10089347;  
 RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
 RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
 awamori.";  
 RL Acta Crystallogr. D 55:403-407 (1999).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=99192677; PubMed=10089508;  
 RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
 RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
 and analysis of ligand-induced conformational change.";  
 RL Acta Crystallogr. D 54:1319-1335 (1998).  
 RN [19]  
 RP CHARACTERIZATION OF LACTOFERROXINS.  
 RX MEDLINE=91166929; PubMed=1369293;  
 RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
 RT "Isolation and characterization of opioic antagonist peptides derived  
 from human lactoferrin.";  
 RL Agric. Biol. Chem. 54:1803-1810 (1990).  
 RN [20]  
 RP VARIANTS THR-30 AND ARG-48.  
 RX PubMed=9873069;  
 RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
 RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
 RA Sugar J., Kumaranickavel G., Munier F., Schorderet D.F.,  
 RA El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
 RA Hejmancik J.F., Teng C.T.;  
 RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
 corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
 RL Mol. Vision 4:31-32 (1998).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
 ACTIVITY. LACTOFERRIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
 LACTOFERRIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
 FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X53961; CAA37914.1; -;  
 DR EMBL; U07643; AAB60324.1; -;  
 DR EMBL; M93150; AAA36159.1; -;  
 DR EMBL; M83202; AAA59511.1; -;  
 DR EMBL; M83205; AAA58656.1; -;  
 DR EMBL; M18642; AAA86655.1; -;  
 DR EMBL; AF332168; AAC48753.1; -;  
 DR EMBL; BC015822; AAH15822.1; -;  
 DR EMBL; BC015823; AAH15823.1; -;  
 DR EMBL; M73700; AAA59479.1; -;  
 DR EMBL; X52941; CAA37116.1; -;  
 DR EMBL; U95626; AAB57795.1; -;  
 DR PIR; S11228; TFHUL.  
 DR PDB; 1LCP; 31-AUG-94.  
 DR PDB; 1LCT; 31-OCT-93.  
 DR PDB; 1LFG; 31-JUL-94.  
 DR PDB; 1LPH; 31-OCT-93.  
 DR PDB; 1LFT; 31-OCT-93.  
 DR PDB; 1LGG; 31-AUG-94.  
 DR PDB; 1LGC; 31-AUG-94.  
 DR PDB; 1BKA; 08-NOV-96.  
 DR PDB; 1DSN; 08-MAR-96.  
 DR PDB; 1HSE; 12-MAR-97.  
 DR PDB; 1VPD; 21-APR-97.

Query Match 95.8%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00024;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKNRKR 12  
 |||||:|||||  
 Db 39 CFQWQKNRKR 50

## RESULT 2

TRFL CAMDR  
 ID TRFL CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUMG; Q9MZG5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 ON NCBI\_TaxID=9838;  
 RN [1]\_TaxID=9838;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AJ131674; CAB53387.1; -;  
 EMBL; AF165879; AAF82241.1; -;  
 HSP; O77811; 1B1X.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; Transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.  
 FT SIGNAL. 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 F -> S (IN REF. 2).  
 FT CONFLICT 261 261 G -> A (IN REF. 2).  
 FT CONFLICT 304 304 S -> P (IN REF. 2).  
 FT CONFLICT 330 330 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 492 494 L -> F (IN REF. 2).  
 FT CONFLICT 506 506 A -> P (IN REF. 2).  
 FT CONFLICT 609 609 R -> Q (IN REF. 2).  
 FT CONFLICT 642 642  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 71.8%; Score 51; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.21;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQKNRKR 12  
 |||||:|||||  
 Db 38 CAQWQKNRKR 49

## RESULT 3

TRFL CAPHI  
 ID TRFL CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 ON NCBI\_TaxID=9925;  
 RN [1]\_TaxID=9925;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Nocart M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine IL2 syntenic group".  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332 (1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.





FT REPEAT 358 707  
 FT DISULFID 27 63  
 FT BY SIMILARITY.  
 FT DISULFID 37 54  
 FT BY SIMILARITY.  
 FT DISULFID 133 216  
 FT BY SIMILARITY.  
 FT DISULFID 175 191  
 FT BY SIMILARITY.  
 FT DISULFID 188 199  
 FT BY SIMILARITY.  
 FT DISULFID 249 263  
 FT BY SIMILARITY.  
 FT DISULFID 366 398  
 FT BY SIMILARITY.  
 FT DISULFID 376 389  
 FT BY SIMILARITY.  
 FT DISULFID 423 702  
 FT BY SIMILARITY.  
 FT DISULFID 443 665  
 FT BY SIMILARITY.  
 FT DISULFID 475 550  
 FT BY SIMILARITY.  
 FT DISULFID 499 693  
 FT BY SIMILARITY.  
 FT DISULFID 509 523  
 FT BY SIMILARITY.  
 FT DISULFID 520 533  
 FT BY SIMILARITY.  
 FT DISULFID 591 605  
 FT BY SIMILARITY.  
 FT DISULFID 643 648  
 FT BY SIMILARITY.  
 FT METAL 78 78  
 FT IRON 1 (BY SIMILARITY).  
 FT METAL 110 110  
 FT IRON 1 (BY SIMILARITY).  
 FT METAL 210 210  
 FT IRON 1 (BY SIMILARITY).  
 FT METAL 271 271  
 FT IRON 1 (BY SIMILARITY).  
 FT METAL 413 413  
 FT IRON 2 (BY SIMILARITY).  
 FT METAL 451 451  
 FT IRON 2 (BY SIMILARITY).  
 FT METAL 544 544  
 FT IRON 2 (BY SIMILARITY).  
 FT METAL 613 613  
 FT IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139  
 FT ANION (POTENTIAL).  
 FT BINDING 481 481  
 FT ANION (POTENTIAL).  
 FT CARBOHYD 118 118  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2  
 FT MR -> IQG (IN REF. 1).  
 FT CONFLICT 25 25  
 FT R -> Q (IN REF. 2).  
 FT CONFLICT 82 82  
 FT M -> L (IN REF. 2).  
 FT CONFLICT 359 359  
 FT S -> T (IN REF. 2).  
 FT CONFLICT 382 382  
 FT A -> D (IN REF. 1).  
 FT CONFLICT 449 449  
 FT E -> G (IN REF. 2).  
 FT CONFLICT 629 629  
 FT L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77865 MW; F26AE03404C19A8 CRC64;

Query Match 63.4%; Score 45; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 2.3;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKMKRV 11  
 DB 37 CLRQWQMKRV 47

RESULT 5  
 ID DYHB\_CHLRE STANDARD; PRT; 4568 AA.  
 AC Q39565;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Dynein beta chain, flagellar outer arm.  
 GN ODA4 OR ODA-4 OR SUP1.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]\_TaxID=3055;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=219T;  
 RX MEDLINE=94274778; PubMed=8006077;  
 RA Mitchell D.R., Brown K.S.;  
 RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes."  
 RL J. Cell Sci. 107:635-644 (1994).  
 CC -!- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.  
 CC DYNEIN HAS ATPASE ACTIVITY.  
 CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.

CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U02963; AAA19956.1; --  
 CC InterPro; IPR004273; Dynein\_heavy.  
 CC Pfam; PF03028; Dynein\_heavy; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 KW Coiled coil.  
 FT DOMAIN 277 293 COILED COIL (POTENTIAL).  
 FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).  
 FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).  
 FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).  
 FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).  
 FT DOMAIN 2017 2045 COILED COIL (POTENTIAL).  
 FT DOMAIN 2831 2848 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).  
 FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).  
 FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1919 1926 ATP (POTENTIAL).  
 FT NP\_BIND 2202 2209 ATP (POTENTIAL).  
 FT NP\_BIND 2530 2537 ATP (POTENTIAL).  
 FT NP\_BIND 2879 2886 ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;  
 Query Match 63.4%; Score 45; DB 1; Length 4568;  
 Best Local Similarity 50.0%; Pred. No. 15;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFQWQKMKRV 12  
 DB 1852 CFQWQKMKRV 1863  
 RESULT 6  
 ID TRFL\_HORSE STANDARD; PRT; 695 AA.  
 AC Q79811;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin) (Fragment).  
 GN LTF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "cDNA sequence of mare lactoferrin."  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RC TISSUE=Milk;  
 RX MEDLINE=99296631; PubMed=10366507;  
 RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A resolution."  
 RL J. Mol. Biol. 289:303-317 (1999).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.



```

FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).
SQ SEQUENCE 1135 AA; 126315 MW; AFCC934B7592DE4D CRC64;

Query Match
Best Local Similarity 54.9%; Score 39; DB 1; Length 1135;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKNRKY 11
DB 775 CLEWKNMQKI 785

RESULT 8
RL28 THEN
ID RL28 THEN STANDARD; PRT; 62 AA.
AC Q8R9UL;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 50S ribosomal protein L28.
GN RPB OR TTE1495.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
ON NCBI_TaxID=119072;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:1689-700(2002).
CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE013107; AAM24713.1; -
CC TIGR; VC2181; -
CC InterPro; IPR004565; LoLB.
CC Pfam; PF03550; LoLB; 1.
CC Chaperone; Outer membrane; Lipoprotein; Transport; Protein transport;
KW Signal; Complete proteome.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 211 OUTER-MEMBRANE LIPOPROTEIN LoLB.
FT LIPID 27 21 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE 211 AA; 24379 MW; F1EF70858484177E CRC64;

Query Match
Best Local Similarity 53.5%; Score 38; DB 1; Length 211;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWQKNRKY 11
DB 71 FQWQKSPQKL 80

RESULT 10
SYH CHLMU STANDARD; PRT; 428 AA.
AC Q9PUJ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE (HISRS).
GN HISS OR TCO830.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Base S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
CC -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF002349; AAP39630.1; -  
 CC HSSP; O32422; 1QE0.  
 CC TIGR; TC0830; -  
 CC InterPro; IPR002106; AATRNA\_ligaseII.  
 CC InterPro; IPR004154; HGTP\_anticonodon.  
 CC InterPro; IPR004516; Hiss.  
 CC InterPro; IPR002314; tRNA-synt\_2b.  
 CC Pfam; PF00587; tRNA-synt\_2b; 1.  
 CC Pfam; PF03129; HGTP\_anticonodon; 1.  
 CC TIGRFAMs; TIGR00442; hiss; 1.  
 CC PROSITE; PS0862; AA\_TRNA\_LIGASE II; 1.  
 CC Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 CC Complete proteome.  
 KW  
 SQ SEQUENCE 428 AA; 48939 MW; 9CF859ED0E689DDF CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 428;  
 Best Local Similarity 55.6%; Pred.No. 23;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFQWQKNMR 9  
 Db 348 CFQWQKNMR 356  
 ID NQRC CHLTR STANDARD; PRT; 316 AA.  
 AC O84281;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable Na(+)-translocating NADH-quinone reductase subunit C  
 DE (EC 1.6.5.-) (Na(+)-translocating NQR subunit C) (Na(+)-NQR subunit C)  
 DE (NQR complex subunit C) (NQR-1 subunit C).  
 GN NQR OR NQR3 OR CT279.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UN-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis";  
 RL Science 282:754-759(1998).  
 CC -1- FUNCTION: NQR COMPLEX CATALYZES THE REDUCTION OF UBIQUINONE-1 TO  
 CC UBIQUINOL BY TWO SUCCESSIVE REACTIONS, COUPLED WITH THE TRANSPORT  
 CC OF NA(+) IONS FROM THE CYTOPLASM TO THE PERIPLASM. NQRA TO NQRE  
 CC ARE PROBABLY INVOLVED IN THE SECOND STEP, THE CONVERSION OF  
 CC UBISEMQUINONE TO UBIQUINOL (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone + Na(+) (In) = NAD(+) +  
 CC ubiquinol + Na(+) (Out).  
 CC -1- COFACTOR: FMN (BY SIMILARITY).  
 CC -1- SUBUNIT: COMPOSED OF SIX SUBUNITS; NQRA, NQRB, NQRC, NQRD, NQRE  
 CC AND NQRF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Inner membrane (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE NQRC FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF001300; AAC67872.1; -  
 CC KW Oxidoreductase; NAD; Ubiquinone; Transport; Sodium transport;  
 CC Flavoprotein; FMN; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 13 35 POTENTIAL.  
 FT BINDING 280 280 FMN (BY SIMILARITY).  
 SQ SEQUENCE 316 AA; 34412 MW; 4AA9442245DE6250 CRC64;  
 Query Match 52.1%; Score 37; DB 1; Length 316;  
 Best Local Similarity 85.7%; Pred.No. 25;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QWQKNMR 9  
 Db 224 QWQKNMR 230  
 ID CATL MOUSE STANDARD; PRT; 334 AA.  
 AC P06797;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP).  
 GN CTSL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=88076849; PubMed=3689328;  
 RA Troen B.R., Gal S., Gottesman M.M.;  
 RT "Sequence and expression of the cDNA for MEP (major excreted  
 RT protein), a transformation-regulated secreted cathepsin";  
 RL Biochem. J. 246:731-735(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8213715; PubMed=2835398;  
 RA Joseph L.J., Chang L.C., Stamenkovich D., Sukhatme V.P.;  
 RT "Complete nucleotide and deduced amino acid sequences of human and  
 RT murine preprocathepsin L. An abundant transcript induced by  
 RT transformation of fibroblasts";  
 RL J. Clin. Invest. 81:1621-1629(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87033683; PubMed=3533924;  
 RA Portnoy D.A., Erickson A.H., Kochan J., Ravetch J.V., Unkeless J.C.;  
 RT "Cloning and characterization of a mouse cysteine proteinase";  
 RL J. Biol. Chem. 261:14697-14703(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND CARBOHYDRATE-LINKAGE SITE.  
 RC TISSUE=Liver;  
 RX MEDLINE=91112761; PubMed=2275556;  
 RA Stearns N.A., Dong J., Pan J.X., Brenner D.A., Sahagian G.G.;  
 RT "Comparison of cathepsin L synthesized by normal and transformed  
 RT cells at the gene, message, protein, and oligosaccharide levels";  
 RL Arch. Biochem. Biophys. 283:447-457(1990).  
 RN [5]  
 RP SEQUENCE OF 89-300 FROM N.A.  
 RC STRAIN=BNL;  
 RX MEDLINE=86271744; PubMed=3755373;  
 RA Denhardt D.T., Hamilton R.T., Parfett C.L.J., Edwards D.R.,  
 RA Pierre R.S., Waterhouse P., Nilsson-Hamilton M.;  
 RT "Close relationship of the major excreted protein of transformed  
 RT murine fibroblasts to thiol-dependent cathepsins";

Cancer Res. 46:4590-4593 (1986).

-1- FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN LYSOSOMES.

-1- CATALYTIC ACTIVITY: Specificity close to that of papain. As compared to cathepsin B, cathepsin L exhibits higher activity towards protein substrates, but has little activity on Z-Arg-Arg-NHMe, and no peptidyl-diesterase activity.

-1- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE BONDS.

-1- SUBCELLULAR LOCATION: Lysosomal.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X06086; CAA29470.1; -;  
 EMBL; J02583; AAA37445.1; -;  
 EMBL; M20495; AAA39984.1; -;  
 EMBL; X04392; CAA27980.1; -;  
 PIR; S01177; KMSL;  
 PIR; S13890; S13890;  
 HSP; P07711; ICLJ.  
 MEROPS; C01.032; -;  
 MGD; MGI:88564; CtsL.  
 InterPro; IPR006688; Peptidase C1.  
 InterPro; IPR000169; SHprot acfite.  
 Pfam; PF00112; Peptidase\_C1; 1.  
 PRINTS; PR00705; PAPAINE.  
 ProDom; PD000158; Peptidase C1; 1.  
 PROSITE; PS00139; THIOL PROTEASE CYS; 1.  
 PROSITE; PS00639; THIOL PROTEASE HIS; 1.  
 PROSITE; PS00640; THIOL PROTEASE ASN; 1.  
 KX Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.  
 FT SIGNAL 1 17  
 FT PROPEP 18 113  
 FT CHAIN 114 298  
 FT CHAIN 291 334  
 FT ACT SITE 138 138  
 FT ACT SITE 276 276  
 FT ACT SITE 300 300  
 FT DISULFID 135 178  
 FT DISULFID 169 211  
 FT DISULFID 269 322  
 FT CARBOHYD 221 221  
 FT CONFLICT 58 58  
 FT CONFLICT 177 177  
 FT SEQUENCE 334 AA; 37547 MW; FE6747043307AD98 CRC64;  
 Query Match 52.1%; Score 37; DB 1; Length 334;  
 Best Local Similarity 55.6%; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WQKNRKR 12  
 Db 52 WEKNRMRIQ 60  
 RESULT 13  
 CATL\_RAT  
 ID CATL\_RAT STANDARD; PRT; 334 AA.  
 AC P07154; Q9QV07;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP)  
 DE (Cyclic protein-2) (CP-2).  
 GN CTSL.  
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Kidney;  
 RX MEDLINE=88030047; PubMed=3666143;  
 RA Ishidoh K., Towatari T., Imajoh S., Kawasaki H., Kominami E.,  
 RA Katunuma N., Suzuki K.;  
 RT "Molecular cloning and sequencing of cDNA for rat cathepsin L";  
 RL FEBS Lett. 223:69-73(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90092543; PubMed=2599113;  
 RX Ishidoh K., Kominami E., Suzuki K., Katunuma N.;  
 RT "Gene structure and 5'-upstream sequence of rat cathepsin L";  
 RL FEBS Lett. 259:71-74(1989).  
 RN [3]  
 RP SEQUENCE OF 1-42 FROM N.A.; TISSUE SPECIFICITY, AND INDUCTION.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=21255611; PubMed=11356678;  
 RA Zabludoff S.D., Charron M., DeCervo J.N., Simukova N., Wright W.W.;  
 RT "Male germ cells regulate transcription of the cathepsin L gene by rat Sertoli cells";  
 RL Endocrinology 142:2318-2327(2001).  
 RN [4]  
 RP SEQUENCE OF 88-334 FROM N.A.  
 RC TISSUE=Sertoli cells;  
 RX MEDLINE=92168015; PubMed=1791830;  
 RA Erickson-Lawrence M., Zabludoff S.D., Wright W.W.;  
 RT "Cyclic protein-2, a secretory product of rat Sertoli cells, is the proenzyme form of cathepsin L";  
 RL Mol. Endocrinol. 5:1789-1798(1991).  
 RN [5]  
 RP SEQUENCE OF 18-37, FUNCTION, AND SUBCELLULAR LOCATION.  
 RC STRAIN=Sprague-Dawley; TISSUE=Sertoli cells;  
 RX MEDLINE=95296691; PubMed=7777858;  
 RA Boujrad N., Ogwegbu S.O., Garnier M., Lee C.-H., Martin B.M.,  
 RA Papadopoulos V.;  
 RT "Identification of a stimulator of steroid hormone synthesis isolated from testis";  
 RL Science 268:1609-1612(1995).  
 RN [6]  
 RP SEQUENCE OF 18-28, AND TISSUE SPECIFICITY.  
 RC STRAIN=Sprague-Dawley; TISSUE=Epidermis;  
 RX MEDLINE=20164186; PubMed=10699763;  
 RA Kawada A., Hara K., Kominami E., Tezuka T., Takahashi M., Takahara H.;  
 RT "Precursor of rat epidermal cathepsin L: purification and immunohistochemical localization";  
 RL J. Dermatol. Sci. 23:36-45(2000).  
 RN [7]  
 RP SEQUENCE OF 114-288 AND 291-334.  
 RC TISSUE=Liver;  
 RX MEDLINE=88296890; PubMed=3402618;  
 RA Towatari T., Katunuma N.;  
 RT "Amino acid sequence of rat liver cathepsin L";  
 RL FEBS Lett. 236:57-61(1988).  
 CC -1- FUNCTION: Important for the overall degradation of proteins in lysosomes. Procathepsin L is required for maximal stimulation of sterologenesis by TIMP1.  
 CC -1- CATALYTIC ACTIVITY: Specificity close to that of papain. As compared to cathepsin B, cathepsin L exhibits higher activity towards protein substrates, but has little activity on Z-Arg-Arg-NHMe, and no peptidyl-diesterase activity.  
 CC -1- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE BONDS.  
 CC -1- SUBCELLULAR LOCATION: Procathepsin L is secreted. Mature cathepsin L is lysosomal.  
 CC -1- TISSUE SPECIFICITY: Both mature cathepsin L and procathepsin L are found in the upper epidermis. The lower epidermis predominantly contains procathepsin L. In seminiferous tubules expression is greater at stages VI-VII than at stages IX-XII.  
 CC -1- INDUCTION: Expression in Sertoli cells is repressed by germ cells.

CC use by non-profit institutions as long as its content is in no way  
CC the European Bioinformatics Institute. There are no restrictions on its

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AL132798; CAB60697.1; -  
 DR InterPro; IPR005289; GTP-binding\_dom.  
 DR InterPro; IPR002917; MMR\_HSR1.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR InterPro; IPR004520; ThdF.  
 DR Pfam; PF01926; MMR\_HSR1; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR TIGRFAMs; TIGR00450; thdF\_1.  
 DR TIGRFAMs; TIGR00650; MG442; 1.  
 KW tRNA processing; Mitochondrion; GTP-binding; Transit peptide.  
 FT TRANSIT ?  
 FT CHAIN ? 496 MITOCHONDRION (POTENTIAL).  
 FT NP BIND 246 253 GTPASE MSS1 HOMOLOG.  
 FT NP BIND 246 253 GTP (POTENTIAL).  
 SQ SEQUENCE 496 AA; 55315 MW; 092F7E1E1CDD80D2 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 496;  
 Best Local Similarity 41.7%; Pred. No. 40;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQKMKRKYR 12  
 |||:|:|:|  
 Db 178 CFWRKKLIEYR 189

Search completed: February 21, 2003, 07:27:57  
 Job time : 6.6 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107B-81  
Perfect score: 71  
Sequence: 1 CFQWQKMKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL 21:.\*
- 1: sp\_archea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	63	88.7	711	4 Q8TCD2	Q8tcd2 homo sapien
2	59	83.1	38	4 Q9UCY5	Q9ucy5 homo sapien
3	54	76.1	33	6 Q9TR80	Q9tr80 ovis aries
4	43	60.6	279	16 Q8XSE2	Q8xse2 ralstonia s
5	43	60.6	469	9 Q38115	Q38115 bacterioph
6	43	60.6	681	5 Q9XVD1	Q9xvd1 caenorhabdi
7	42	59.2	511	16 Q8Z462	Q8z462 salmonella
8	42	59.2	584	17 Q9HPA3	Q9hpa3 halobacteri
9	41	57.7	275	5 Q93780	Q93780 caenorhabdi
10	41	57.7	570	10 Q8S487	Q8s487 zea mays (m
11	40	56.3	115	15 Q9Q9L0	Q9q9l0 human immun
12	40	56.3	121	10 Q9L7N4	Q9l7n4 arabidopsis
13	40	56.3	306	4 Q8TAX2	Q8tax2 homo sapien
14	40	56.3	466	4 Q9NUS2	Q9nus2 homo sapien
15	39	54.9	274	4 Q96M21	Q96m21 homo sapien
16	39	54.9	283	16 Q9WYQ1	Q9wyq1 thermotoga

17	39	54.9	298	16 Q8YP77	Q8yp77 anabaena sp
18	39	54.9	589	16 Q8RGT4	Q8rgt4 fusobacteri
19	39	54.9	602	10 Q8S013	Q8s013 oryza sativ
20	39	54.9	1274	10 Q9S287	Q9s287 arabidopsis
21	38	53.5	62	16 Q8R9U1	Q8r9ul thermococci
22	38	53.5	148	10 Q9XHF1	Q9xhpl sesamum ind
23	38	53.5	205	2 Q9APZ3	Q9apz3 vibrio chol
24	38	53.5	205	16 Q986A0	Q986a0 rhizobium l
25	38	53.5	206	16 Q8UHC2	Q8uhc2 agrobacteri
26	38	53.5	206	16 Q92RH8	Q92rh8 rhizobium m
27	38	53.5	208	16 Q8YFK3	Q8yfk3 brucella me
28	38	53.5	282	16 Q98Q19	Q98q19 mycoplasma
29	38	53.5	289	10 Q9C6N2	Q9c6n2 arabidopsis
30	38	53.5	393	16 Q9ZKP4	Q9zkp4 helicobacte
31	38	53.5	443	16 Q97TQ8	Q97tq8 clostridium
32	38	53.5	459	4 Q9NZW0	Q9nzw0 homo sapien
33	38	53.5	460	4 Q9NZW3	Q9nzw3 homo sapien
34	38	53.5	632	4 Q94937	Q94937 homo sapien
35	38	53.5	759	16 Q9Z955	Q9z955 chlamydia p
36	38	53.5	866	10 Q9FHI9	Q9fhi9 arabidopsis
37	38	53.5	932	5 Q19153	Q19153 caenorhabdi
38	38	53.5	1432	10 Q23230	Q23230 arabidopsis
39	37	52.1	81	15 Q90863	Q90863 human immun
40	37	52.1	81	15 Q90884	Q90884 human immun
41	37	52.1	99	16 Q92GL6	Q92gl6 rickettsia
42	37	52.1	108	9 Q8SC55	Q8sc55 stx2 conver
43	37	52.1	112	17 Q971N7	Q971n7 sulfobolus
44	37	52.1	122	15 Q72904	Q72904 human immun
45	37	52.1	248	2 Q8RMB8	Q8rmb8 cytophaga j

ALIGNMENTS

RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.

AC Q8TCD2; (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE Lactotransferrin.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;

Q8TCD2 [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PROSTATE;

RA Strausberg R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC022347; AAH22347.1; -.

SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 88.7%; Score 63; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0041;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQKMKRV 11  
| | | | |  
Db 39 CFQWQKMKRV 49

RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.

AC Q9UCY5; (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE Lactoferrin homolog (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 RN NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 seminal plasma";  
 RL Jpn. J. Legal Med. 49:281-293(1995).  
 DR HSP: P02788; IBA.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDEB CRC64;

Query Match 83.1%; Score 59; DB 4; Length 38;  
 Best Local Similarity 90.9%; Pred. No. 0.001; Indels 0;  
 Matches 10; Conservative 1; Mismatches 0; Gaps 0;

QY 2 FQWQKMKVR 12  
 [1]  
 Db 21 FQWQKMKVR 31

RESULT 3  
 Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.-Y., Jolles P., Migliore-Samour D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32(1995).  
 DR HSP: O77698; ICE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 76.1%; Score 54; DB 6; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.007;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQWQKMKVR 11  
 [1]  
 Db 19 FQWQKMKVR 29

RESULT 4  
 Q8XSE2 PRELIMINARY; PRT; 279 AA.  
 AC Q8XSE2;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DE Putative ICC protein homolog.  
 GN ICC OR RSP0534 OR RSP0414.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OG Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;

RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Signier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646079; CAD17685.1;  
 DR InterPro: IPR004843; M-peptidase.  
 DR Pfam: PF00149; Metallophos; 1.  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 279 AA; 31541 MW; ABH38818004B2EDA CRC64;

Query Match 60.6%; Score 43; DB 16; Length 279;  
 Best Local Similarity 50.0%; Pred. No. 6.6;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQKMKVR 12  
 [1]  
 Db 244 CFQWQKMKVR 255

RESULT 5  
 Q38115 PRELIMINARY; PRT; 469 AA.  
 AC Q38115;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE ORF29.  
 OS Bacteriophage rlt.  
 OC Viruses.  
 OX NCBI\_TaxID=43685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96332668; PubMed=8730874;  
 RA Nauta A., Van Sinderen D., Karsens H., Smit E., Venema G., Kok J.;  
 RT "Inducible gene expression mediated by a repressor-operator system  
 isolated from Lactococcus lactis bacteriophage rlt.";  
 RL Mol. Microbiol. 19:1331-1341(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96332669; PubMed=8730875;  
 RA Van Sinderen D., Karsens H., Kok J., Terpstra P., Ruiter M.H.,  
 RA Venema G., Nauta A.;  
 RT "Sequence analysis and molecular characterization of the temperate  
 lactococcal bacteriophage rlt.";  
 RL Mol. Microbiol. 19:1343-1355(1996).  
 DR EMBL: U38906; AAB18704.1;  
 SQ SEQUENCE 469 AA; 53160 MW; 1F8E02D4325C6BB9 CRC64;

Query Match 60.6%; Score 43; DB 9; Length 469;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQKMKVR 10  
 [1]  
 Db 39 CYPWQKMKVR 48

RESULT 6  
 Q9XVD1 PRELIMINARY; PRT; 681 AA.  
 AC Q9XVD1;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE C24H11.8 protein.  
 GN C24H11.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
  investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z81475; CAB03914.1; -.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR000636; M+channel_ilg.
DR Pfam; PF00520; ion trans; 1.
SQ SEQUENCE 681 AA; 78178 MW; D36AC05C3FA029CE CRC64;

Query Match 50.6%; Score 43; DB 5; Length 681;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQKNMKVR 12
   ||||| :|||
Db 192 RWQKNRRVR 201

RESULT 7
Q82462
ID Q82462 PRELIMINARY; PRT; 511 AA.
AC Q82462;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein STY3070.
GN STY3070.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL: AL627276; CAD06049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DDD124E10D178B CRC64;

Query Match 59.2%; Score 42; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWKNMKVR 12
   ||||| :|||
Db 350 CFQWDMNKAQR 361

RESULT 8
Q9HPA3
ID Q9HPA3 PRELIMINARY; PRT; 584 AA.
AC Q9HPA3;

```

```

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE VNG1732C.
GN VNG1732C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Oher A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005078; AAG19967.1; -.
DR InterPro; IPR001646; Spectide repeat.
DR InterPro; IPR001622; K+channel_pore.
DR Pfam; PF00805; Pentapeptide; 2.
KW Complete proteome.
SQ SEQUENCE 584 AA; 65151 MW; 21BF5DF0486CCC6 CRC64;

Query Match 59.2%; Score 42; DB 17; Length 584;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNMKVR 12
   ||||| :|||
Db 445 CFQWKNMKVR 456

RESULT 9
Q93780
ID Q93780 PRELIMINARY; PRT; 275 AA.
AC Q93780;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F53H4.4 protein.
GN F53H4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiidae; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Dobson R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
  investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z81089; CAB03137.1; -.
SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 57.7%; Score 41; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWKNMKVR 12
   ||||| :|||
Db 262 FQWKNMKVR 272

```

CDNA FLJ11175 fis, clone PLACE1007375, weakly similar to phorbol

Search completed: February 21, 2003, 07:44:36  
Job time : 21.8 secs

```

DE ester/diacylglycerol-binding protein UNC-13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002037; BAA92048.1; -.
DR HSP; P21707; IBIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00168; C2; 1.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM02239; C2; 1.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 466 AA; 53192 MW; E4113A5062F58D6E CRC64;

Query Match          56.3%; Score 40; DB 4; Length 466;
Best Local Similarity 55.8%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQKNMR 9
Db |||||:
269 CFQWQKNMR 277

RESULT 15
Q96M21 ID Q96M21 PRELIMINARY; PRT; 274 AA.
AC Q96M21;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDNA FLJ32891 fis, clone TEST12004929.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057453; BAB71493.1; -.
SQ SEQUENCE 274 AA; 30083 MW; 1DD43654D4135B2F CRC64;

Query Match          54.9%; Score 39; DB 4; Length 274;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQKNMRKVR 12
Db |||||:
66 CFQWQKNMRKVR 77

```

GenCore version 5.1.3

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 Seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107b-82

Perfect score: 70

Sequence: 1 CFQQRNLRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	12	21	AAV78082
2	68	97.1	12	21	AAV78038
3	68	97.1	12	21	AAV78046
4	68	97.1	12	21	AAV78047
5	68	97.1	13	21	AAV78037
6	68	97.1	13	21	AAV78048
7	68	97.1	13	21	AAV78049
8	68	97.1	14	21	AAV78036
9	68	97.1	14	21	AAV78050
10	68	97.1	14	21	AAV78051
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin
					Human lactoferrin

11	68	97.1	15	17	AAV98554	Peptide for anti-u
12	68	97.1	15	21	AAV78035	Human lactoferrin
13	68	97.1	15	21	AAV78062	Human lactoferrin
14	68	97.1	15	21	AAV78063	Human lactoferrin
15	68	97.1	16	21	AAV78031	Human lactoferrin
16	68	97.1	16	21	AAV78064	Human lactoferrin
17	68	97.1	17	21	AAV78065	Human lactoferrin
18	68	97.1	17	21	AAV78034	Human lactoferrin
19	68	97.1	17	21	AAV78066	Human lactoferrin
20	68	97.1	17	21	AAV78067	Human lactoferrin
21	68	97.1	18	15	AAV69352	Human lactoferrin
22	68	97.1	18	17	AAV13397	Human lactoferrin
23	68	97.1	18	21	AAV78033	Human lactoferrin
24	68	97.1	19	21	AAV68867	Human lactoferrin
25	68	97.1	19	21	AAV78032	Amino acid sequenc
26	68	97.1	20	13	AAV21810	Human lactoferrin
27	68	97.1	20	14	AAV44841	Anti microbial pep
28	68	97.1	20	15	AAV48530	Lactoferrin-relate
29	68	97.1	20	15	AAV48531	Lactoferrin-derive
30	68	97.1	20	15	AAV57461	Lactoferrin-derive
31	68	97.1	20	15	AAV57462	Lactoferrin-derive
32	68	97.1	20	16	AAV84698	Lactoferrin-derive
33	68	97.1	20	16	AAV84699	Bovine lactoferrin
34	68	97.1	20	16	AAV80263	Bovine lactoferrin
35	68	97.1	20	16	AAV80264	Anti-parasitic lac
36	68	97.1	20	17	AAV98553	Anti-parasitic lac
37	68	97.1	20	17	AAV91852	Peptide for anti-u
38	68	97.1	20	17	AAV03045	Lactoferrin-derive
39	68	97.1	20	17	AAV90607	Lactoferrin-derive
40	68	97.1	20	17	AAV87621	Lactoferrin-derive
41	68	97.1	20	17	AAV87622	Lactoferrin-derive
42	68	97.1	20	18	AAV26150	Lactoferrin-derive
43	68	97.1	20	18	AAV14036	Lactoferrin deriva
44	68	97.1	20	19	AAV70310	Anti-parasitic pep
45	68	97.1	20	19	AAV53224	Thrombus formation
						Lactoferrin hydrol

## ALIGNMENTS

RESULT 1  
AAV78082  
ID AAV78082 standard; Peptide; 12 AA.  
XX AAV78082;  
AC AAV78082;  
DT 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:82.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI WPI; 2000-147388/13.  
XX

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 22; Page 36; 102pp; English.  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 XX Query Match 100.0%; Score 70; DB 21; Length 12;  
 XX Best Local Similarity 100.0%; Pred. No. 1.7e-05; Indels 0; Gaps 0;  
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRLNRKVR 12  
 DB 1 CFQWQRLNRKVR 12  
 RESULT 2  
 AAY78038  
 ID AAY78038 standard; Peptide; 12 AA.  
 XX AAY78038;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:38.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 70; 102pp; English.  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 XX Query Match 97.1%; Score 68; DB 21; Length 12;  
 XX Best Local Similarity 91.7%; Pred. No. 3.7e-05; Indels 0; Gaps 0;  
 XX Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRLNRKVR 12  
 DB 1 CFQWQRLNRKVR 12  
 RESULT 3  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 XX AAY78046;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:46.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 15; Page 35; 102pp; English.  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.1%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 3.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKVR 12  
 DB 1 CFQWQRMNRKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:47.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumours. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.1%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 3.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMNRKVR 12  
 DB 1 CFQWQRMNRKVR 12

RESULT 5  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:37.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumours. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.

XX SQ Sequence 13 AA;

Query Match 97.1%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 4e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12  
 DB 2 CFQWQNRNRKVR 13

## RESULT 6

AAV78048  
 ID AAY78048 standard; Peptide; 13 AA.

XX AAY78048;  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

XX inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 97.1%; Score 68; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 4e-05; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12

DB 2 CFQWQNRNRKVR 13

## RESULT 7

AAV78049

ID AAY78049 standard; Peptide; 13 AA.

XX AAY78049;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

XX inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 97.1%; Score 68; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 4e-05; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12

DB 2 CFQWQNRNRKVR 13

## RESULT 8

AAV78036

ID AAY78036 standard; Peptide; 14 AA.

XX AAY78036;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.



XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 69; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 97.1%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 4.4e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQNRNRKVR 12  
 Db |||||:||||  
 3 CFQWQNRNRKVR 14  
 RESULT 9  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX  
 AC AAY78050;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:50.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 15; Page 75; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 97.1%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 4.4e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQNRNRKVR 12  
 Db |||||:||||  
 3 CFQWQNRNRKVR 14  
 RESULT 10  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX  
 AC AAY78051;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:51.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 XX Claim 18; Page 75; 102pp; English.  
 PS  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, colitis, and Candida infection on a mucosal  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 97.1%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 4.4e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRLRKVR 12  
 Db 3 CFQWQRLRKVR 14  
 RESULT 11  
 AAR98554  
 ID AAR98554 standard; Peptide; 15 AA.  
 XX  
 AC AAR98554;  
 XX  
 DT 12-NOV-1996 (first entry)  
 XX  
 DE Peptide for anti-ulcer agent.  
 XX  
 KW anti-ulcer agent; low toxicity; stable; heat-resistant.  
 XX  
 OS Synthetic.  
 XX  
 XX JP08143468-A.  
 PN  
 XX 04-JUN-1996.  
 PD  
 XX 17-NOV-1994; 94JP-0283869.  
 PF  
 XX 17-NOV-1994; 94JP-0283869.  
 PR  
 XX (MORG ) MORINAGA MILK IND CO LTD.  
 PA  
 XX WPI; 1996-318857/32.  
 DR  
 XX Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble  
 PT  
 XX Claim 1; Page 11; 11pp; Japanese.  
 PS

XX AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.  
 XX  
 SQ Sequence 15 AA;  
 Query Match 97.1%; Score 68; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRLRKVR 12  
 Db 2 CFQWQRLRKVR 13  
 RESULT 12  
 AAY78035  
 ID AAY78035 standard; Peptide; 15 AA.  
 XX  
 AC AAY78035;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:35.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 XX Claim 12; Page 69; 102pp; English.  
 PS  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, colitis, and Candida infection on a mucosal  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 15 AA;

Query Match 97.1%; Score 68; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNLRKVR 12  
| | | | | : | | | | |  
DB 4 CFQWQRNLRKVR 15

RESULT 13  
AAV78062  
ID AAV78062 standard; Peptide; 15 AA.

XX AC AAV78062;  
XX DT 25-APR-2000 (first entry)  
XX DE Human lactoferrin derived peptide SEQ ID NO:62.  
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.

XX OS Homo sapiens.  
XX OS Synthetic.  
XX PN WO200001730-A1.  
XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
XX PR 06-JUL-1998; 98SE-0002441.  
XX PR 17-JUL-1998; 98SE-0002562.  
XX PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -  
XX PS Claim 15; Page 81; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
XX CC lactoferrin. The peptides are taken up in the intestine through  
XX CC binding to specific lactoferrin receptors and are then transported  
XX CC through the circulation. A medicinal product of the peptide or fragment  
XX CC can be used for treating and/or prevention of infections (such as  
XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
XX CC membrane), inflammations and/or tumours. The peptides can also be used  
XX CC in food stuffs such as infant formula food. The peptides are also  
XX CC fungicidal and bactericidal and may also be used as preservatives.  
XX CC Even though native human lactoferrin have been shown to have desired  
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
XX CC cannot be used clinically on a broad basis because of high production  
XX CC costs. Therefore, provision of peptides based on lactoferrin would  
XX CC enable them to be used for the same purposes as lactoferrin at lower  
XX CC cost.

XX SQ Sequence 15 AA;

Query Match 97.1%; Score 68; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNLRKVR 12  
| | | | | : | | | | |  
DB 4 CFQWQRNLRKVR 15

RESULT 14  
AAV78063

ID AAV78063 standard; Peptide; 15 AA.

XX AC AAV78063;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:63.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.

XX OS Homo sapiens.  
XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -  
XX PS Claim 18; Page 81; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
XX CC lactoferrin. The peptides are taken up in the intestine through  
XX CC binding to specific lactoferrin receptors and are then transported  
XX CC through the circulation. A medicinal product of the peptide or fragment  
XX CC can be used for treating and/or prevention of infections (such as  
XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
XX CC membrane), inflammations and/or tumours. The peptides can also be used  
XX CC in food stuffs such as infant formula food. The peptides are also  
XX CC fungicidal and bactericidal and may also be used as preservatives.  
XX CC Even though native human lactoferrin have been shown to have desired  
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
XX CC cannot be used clinically on a broad basis because of high production  
XX CC costs. Therefore, provision of peptides based on lactoferrin would  
XX CC enable them to be used for the same purposes as lactoferrin at lower  
XX CC cost.

XX SQ Sequence 15 AA;

Query Match 97.1%; Score 68; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNLRKVR 12  
| | | | | : | | | | |  
DB 4 CFQWQRNLRKVR 15

RESULT 15

AAV78031

ID AAV78031 standard; Peptide; 16 AA.

XX AC AAV78031;

```

XX 25-APR-2000 (first entry)
DT
XX
DE Human lactoferrin derived peptide SEQ ID NO:31.
XX
XX Human; lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200001730-A1.
XX
PD 13-JAN-2000.
XX
XX 06-JUL-1999; 99WO-SE01230.
XX
XX 06-JUL-1998; 98SE-0002441.
PR 17-JUL-1998; 98SE-0002562.
PR 29-DEC-1998; 98SE-0004614.
XX
XX (ASCII-) A+ SCI INVEST AB.
XX
XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;
XX
XX WPI; 2000-147388/13.
XX
XX New peptides used for treatment and prevention of infections,
PT inflammations and tumors and for use in infant formula food -
XX
XX Claim 11; Page 68; 102pp; English.
XX
XX AAY78001 to AAY78100 represent peptides having sequences based on human
CC lactoferrin. The peptides are taken up in the intestine through
CC binding to specific lactoferrin receptors and are then transported
CC through the circulation. A medicinal product of the peptide or fragment
CC can be used for treating and/or prevention of infections (such as
CC urinary tract infections, colitis, and Candida infection on a mucosal
CC membrane), inflammations and/or tumors. The peptides can also be used
CC in food stuffs such as infant formula food. The peptides are also
CC fungicidal and bactericidal and may also be used as preservatives.
CC Even though native human lactoferrin have been shown to have desired
CC anti-inflammatory anti-infectious and anti-tumoural properties they
CC cannot be used clinically on a broad basis because of high production
CC costs. Therefore, provision of peptides based on lactoferrin would
CC enable them to be used for the same purposes as lactoferrin at lower
CC cost.
XX
XX Sequence 16 AA;
SQ
Query Match 97.1%; Score 68; DB 21; Length 16;
Best Local Similarity 91.7%; Pred. No. 5e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CFQWRNLRKVR 12
Db 5 CFQWRNLRKVR 16

```

Search completed: February 21, 2003, 07:37:14  
 Job time : 28.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 Seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-82  
Perfect score: 70  
Sequence: 1 CFOWQNRNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	97.1	18	1	US-08-204-487-3
2	68	97.1	18	2	US-08-485-948-8
3	68	97.1	18	2	US-08-628-380-8
4	68	97.1	18	2	US-08-475-055-8
5	68	97.1	20	1	US-07-755-161A-3
6	68	97.1	20	1	US-07-891-174-3
7	68	97.1	20	1	US-08-204-487-1
8	68	97.1	20	1	US-08-256-771-24
9	68	97.1	20	1	US-08-256-771-25
10	68	97.1	20	1	US-08-381-984-24
11	68	97.1	20	1	US-08-381-984-25
12	68	97.1	22	4	US-09-508-734-4
13	68	97.1	24	4	US-09-508-734-6
14	68	97.1	25	1	US-07-755-161A-10
15	68	97.1	25	1	US-07-891-174-10
16	68	97.1	25	1	US-08-204-487-7
17	68	97.1	29	4	US-09-508-734-8
18	68	97.1	36	1	US-07-755-161A-8
19	68	97.1	36	1	US-07-891-174-8
20	68	97.1	36	1	US-08-256-771-30
21	68	97.1	36	1	US-08-381-984-29
22	68	97.1	47	2	US-08-464-182A-6
23	68	97.1	40	2	US-08-406-271-6
24	68	97.1	50	2	US-08-693-274A-7
25	68	97.1	52	4	US-09-017-043A-3
26	68	97.1	53	2	US-08-464-182A-5
27	68	97.1	53	2	US-08-406-271-5

28	68	97.1	54	2	US-08-464-182A-2
29	68	97.1	54	2	US-08-406-271-2
30	68	97.1	694	3	US-08-724-586-2
31	68	97.1	694	4	US-09-421-632-2
32	68	97.1	694	4	US-09-932-190-2
33	68	97.1	705	2	US-08-655-640-2
34	68	97.1	708	2	US-08-655-640-4
35	68	97.1	711	1	US-08-354-019-4
36	68	97.1	711	1	US-08-461-333-4
37	68	97.1	711	3	US-08-464-167-4
38	68	97.1	711	3	US-09-158-313-4
39	68	97.1	711	4	US-08-476-798-4
40	65	92.9	711	1	US-08-145-681-2
41	65	92.9	711	1	US-08-250-308-2
42	65	92.9	711	1	US-08-453-703-2
43	65	92.9	711	2	US-08-456-106-2
44	65	92.9	711	3	US-08-456-108-2
45	65	92.9	711	4	US-09-265-577-2

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5585425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAKOI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,487  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A.  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..18  
OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 4, Appli  
SEQUENCE 4, Appli  
SEQUENCE 4, Appli  
SEQUENCE 4, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli  
SEQUENCE 2, Appli

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 97.1%; Score 68; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12  
Db 1 CFQWQNRNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VLASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,948  
FILING DATE:

## CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995

## APPLICATION NUMBER: 08/418,642

FILING DATE: APRIL 7, 1995

## CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

## INFORMATION FOR SEQ ID NO: 8:

## SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: LF-Cl, 8-25

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 97.1%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12  
Db 1 CFQWQNRNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341

## GENERAL INFORMATION:

APPLICANT: LI YONG MING  
APPLICANT: VLASSARA, HELEN  
APPLICANT: CERAMI, ANTHONY  
TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,380  
FILING DATE: April 4, 1996  
CLASSIFICATION: 514

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 514

## ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

## INFORMATION FOR SEQ ID NO: 8:

## SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: LF-Cl, 8-25

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 97.1%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKVR 12  
Db 1 CFQWQNRNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VLASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;;  
;; COMPUTER READABLE FORM: disk  
;; MEDIUM TYPE: Floppy  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;;  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-C1, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;;  
US-08-475-055-8

Query Match 97.1%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
DB 1 CFQWQNRKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;;  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
;;  
US-07-755-161A-3

Query Match 97.1%; Score 68; DB 1; Length 20;

Best Local Similarity 91.7%; Pred. No. 4.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRNLRKVR 12  
Db 2 CFQWQNRNLRKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 97.1%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRNLRKVR 12  
Db 2 CFQWQNRNLRKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOIKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000



```

; TELEFAX: (517) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
; OTHER INFORMATION: DERIVED FROM HUMAN LACTOFERRIN"
US-08-204-487-1

Query Match 97.1%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.8e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRLRKVR 12
Db 2 CFQWQRLRKVR 13

RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"
US-08-381-984-24

Query Match 97.1%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.8e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRLRKVR 12
Db 2 CFQWQRLRKVR 13

RESULT 9
US-08-256-771-25
; Sequence 25, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are protected to
; OTHER INFORMATION: prevent disulfide bond"
US-08-256-771-25

Query Match 97.1%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.8e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRLRKVR 12
Db 2 CFQWQRLRKVR 13

RESULT 10
US-08-381-984-24
```

; Sequence 24, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
US-08-381-984-24  
Query Match 97.1%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CFQWQRLRKVR 12  
Db 2 CFQWQRLRKVR 13  
RESULT 11  
US-08-381-984-25  
; Sequence 25, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack

; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
US-08-381-984-25  
Query Match 97.1%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CFQWQRLRKVR 12  
Db 2 CFQWQRLRKVR 13  
RESULT 12  
US-09-508-734-4  
; Sequence 4, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; TITLE OF INVENTION: useful microorganism thereof  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KR99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 4

;  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-508-734-4

Query Match 97.1%; Score 68; DB 4; Length 22;  
Best Local Similarity 91.7%; Pred. No. 5.3e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRLRKVR 12  
|||||:||||  
DB 2 CFQWQRLRKVR 13

## RESULT 13

US-09-508-734-6  
; Sequence 6, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KR99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 6  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-508-734-6

Query Match 97.1%; Score 68; DB 4; Length 24;  
Best Local Similarity 91.7%; Pred. No. 5.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRLRKVR 12  
|||||:||||  
DB 3 CFQWQRLRKVR 14

## RESULT 14

US-07-755-161A-10  
; Sequence 10, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/755,161A  
; FILING DATE: 19910905  
; CLASSIFICATION: 530  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-10  
Query Match 97.1%; Score 68; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 6e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWQRLRKVR 12  
|||||:||||  
DB 4 CFQWQRLRKVR 15

RESULT 15  
US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No 5317094  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10  
; Query Match 97.1%; Score 68; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 6e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
; Qy 1 CFQWQWNLKVR 12  
; Db 4 CFQWQWNLKVR 15  
; Search completed: February 21, 2003, 07:50:36  
; Job time : 8.7 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107b-82

Perfect score: 70

Sequence: 1 CFQWRNLRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	97.1	15	9	US-09-798-869-2
2	68	97.1	25	9	US-09-798-869-20
3	68	97.1	694	9	US-10-023-096-2
4	60	85.7	15	9	US-09-798-869-6
5	51	72.9	15	9	US-09-798-869-3
6	51	72.9	25	9	US-09-798-869-23
7	43	61.4	15	9	US-09-798-869-7
8	42	60.0	15	9	US-09-798-869-4
9	42	60.0	25	9	US-09-798-869-22
10	39	55.7	15	9	US-09-798-869-8
11	39	55.7	15	9	US-09-798-869-29
12	39	55.7	15	9	US-09-798-869-30
13	38	54.3	86	9	US-09-798-869-30
14	38	54.3	607	9	US-09-798-869-10
15	38	54.3	688	9	US-09-881-579-15
16	37	52.9	21	10	US-09-864-761-47985
17	37	52.9	846	9	US-10-051-409-4
18	36	51.4	62	10	US-09-815-242-12129
19	36	51.4	62	10	US-09-815-242-13026

20	36	51.4	239	10	US-09-864-761-37353	Sequence 37353, A
21	36	51.4	747	9	US-10-066-500-58	Sequence 58, Appl
22	36	51.4	747	9	US-10-002-796-58	Sequence 58, Appl
23	36	51.4	747	9	US-10-066-273-58	Sequence 58, Appl
24	36	51.4	747	9	US-10-066-494-58	Sequence 58, Appl
25	35	50.0	15	9	US-09-798-869-5	Sequence 5, Appl
26	35	50.0	209	10	US-09-904-536-8	Sequence 8, Appl
27	35	50.0	209	10	US-09-904-536-9	Sequence 9, Appl
28	35	50.0	209	10	US-09-904-536-11	Sequence 11, Appl
29	35	50.0	209	10	US-09-904-536-12	Sequence 12, Appl
30	35	50.0	209	10	US-09-904-536-13	Sequence 13, Appl
31	35	50.0	209	10	US-09-904-536-14	Sequence 14, Appl
32	35	50.0	209	10	US-09-904-536-15	Sequence 15, Appl
33	35	50.0	209	10	US-09-904-536-16	Sequence 16, Appl
34	35	50.0	209	10	US-09-904-536-17	Sequence 17, Appl
35	35	50.0	209	10	US-09-904-536-18	Sequence 18, Appl
36	35	50.0	212	10	US-09-904-536-10	Sequence 10, Appl
37	35	50.0	235	9	US-10-095-449-6	Sequence 6, Appl
38	35	50.0	235	10	US-09-448-378-1	Sequence 1, Appl
39	35	50.0	235	10	US-09-983-806-6	Sequence 6, Appl
40	35	50.0	235	10	US-09-904-536-1	Sequence 1, Appl
41	35	50.0	338	9	US-09-978-295A-119	Sequence 119, App
42	35	50.0	338	9	US-09-978-697-119	Sequence 119, App
43	35	50.0	338	9	US-09-978-192A-119	Sequence 119, App
44	35	50.0	338	9	US-09-999-832A-119	Sequence 119, App
45	35	50.0	338	9	US-09-978-189-119	Sequence 119, App

## ALIGNMENTS

## RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 97.1%; Score 68; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 2.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNLRKVR 12  
| | | | | | | | | | | | | | | |  
DB 3 CFQWRNLRKVR 14

## RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 97.1%; Score 68; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 4.4e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
Db 3 CFQWQNRKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William B.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 97.1%; Score 68; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
Db 22 CFQWQNRKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
US-09-798-869-6

Query Match 85.7%; Score 60; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00054;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
Db 3 CFQWQNRKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 72.9%; Score 51; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.016;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
|:||||:|:  
Db 3 CYQWQRRMKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 72.9%; Score 51; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.026;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
|:||||:|:  
Db 3 CYQWQRRMKL 13

## RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 61.4%; Score 43; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.32;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
|:||||:|:  
Db 3 CYQWQRRMKL 13

## RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 60.0%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
|:||||:|:  
Db 3 CLRQWQNRKV 13

## RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 60.0%; Score 42; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 0.76;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
|:||||:|:  
Db 3 CLRQWQNRKV 13

RESULT 10  
 US-09-798-869-8  
 ; Sequence 8, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
 US-09-798-869-8

Query Match 55.7%; Score 39; DB 9; Length 15;  
 Best Local Similarity 54.5%; Pred. No. 1.4;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
 |||:|:|:  
 Db 3 CLRQWQWEMKLV 13

RESULT 11  
 US-09-798-869-29  
 ; Sequence 29, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 29  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: BOVINE  
 US-09-798-869-29

Query Match 55.7%; Score 39; DB 9; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 1.4;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
 |||:|:|:  
 Db 3 CFRWQWRMKLV 13

RESULT 12  
 US-09-798-869-30

; Sequence 30, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 30  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: BOVINE  
 US-09-798-869-30

Query Match 55.7%; Score 39; DB 9; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 1.4;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRLRKV 11  
 |||:|:|:  
 Db 3 CFRWQWRMKLV 13

RESULT 13  
 US-09-738-626-5715  
 ; Sequence 5715, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKKO  
 ; APPLICANT: SENO, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 5715  
 ; LENGTH: 86  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-5715

Query Match 54.3%; Score 38; DB 9; Length 86;  
 Best Local Similarity 54.5%; Pred. No. 11;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWQRLRKV 12  
 |||:|:|:  
 Db 73 FEYRQLRKIR 83



## RESULT 14

US-09-881-579-10  
; Sequence 10, Application US/09881579  
; Publication No. US20030028904A1  
; GENERAL INFORMATION:  
; APPLICANT: Tina L. Gumienny  
; APPLICANT: Enrico Brugnara  
; APPLICANT: Annie-Carole Tosello-Tramont  
; APPLICANT: Jason M. Kinchen  
; APPLICANT: Michael O. Hengartner  
; APPLICANT: Kodimangalam S. Ravichandran  
; TITLE OF INVENTION: Genes Involved in Engulfment of Dying  
; TITLE OF INVENTION: Cells and Cell Migration  
; FILE REFERENCE: 1314.2008-001  
; CURRENT APPLICATION NUMBER: US/09/881,579  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/285,469  
; PRIOR FILING DATE: 2001-04-19  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 607  
; TYPE: PRI  
; ORGANISM: Homo sapien  
US-09-881-579-10

Query Match 54.3%; Score 38; DB 9; Length 607;  
Best Local Similarity 66.7%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FOWQRLRK 10  
: |||||:  
Db 140 XLWQRLRQ 148

## RESULT 15

US-09-881-579-15  
; Sequence 15, Application US/09881579  
; Publication No. US20030028904A1  
; GENERAL INFORMATION:  
; APPLICANT: Tina L. Gumienny  
; APPLICANT: Enrico Brugnara  
; APPLICANT: Annie-Carole Tosello-Tramont  
; APPLICANT: Jason M. Kinchen  
; APPLICANT: Michael O. Hengartner  
; APPLICANT: Kodimangalam S. Ravichandran  
; TITLE OF INVENTION: Genes Involved in Engulfment of Dying  
; TITLE OF INVENTION: Cells and Cell Migration  
; FILE REFERENCE: 1314.2008-001  
; CURRENT APPLICATION NUMBER: US/09/881,579  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/285,469  
; PRIOR FILING DATE: 2001-04-19  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 688  
; TYPE: PET  
; ORGANISM: Murine  
US-09-881-579-15

Query Match 54.3%; Score 38; DB 9; Length 688;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FOWQRLRK 10  
: |||||:  
Db 220 YLWQRLRQ 228

Search completed: February 21, 2003, 08:08:08  
Job time : 11.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-82

Perfect score: 70

Sequence: 1 CFQQRNLKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	68	97.1	711	1	TFHUL
2	51	72.9	708	2	lactotransferrin p
3	48	68.6	33	2	lactoferrin - goat
4	46	65.7	4568	2	dynein beta heavy
5	42	60.0	298	2	hypothetical prote
6	42	60.0	707	1	lactoferrin precu
7	41	58.6	511	2	hypothetical prote
8	41	58.6	932	2	hypothetical prote
9	40	57.1	205	2	apolipoprotein B-1
10	40	57.1	274	2	hypothetical prote
11	39	55.7	282	2	hypothetical prote
12	39	55.7	681	2	hypothetical prote
13	38	54.3	275	2	hypothetical prote
14	38	54.3	357	2	hypothetical prote
15	38	54.3	535	2	probable cyclochrom
16	38	54.3	543	2	cytochrome P450 ho
17	38	54.3	1432	2	trichohyalin like
18	37	52.9	114	2	carcinoembryonic a
19	37	52.9	121	2	hypothetical prote
20	37	52.9	222	2	hypothetical prote
21	37	52.9	289	2	33.3K hypothetical
22	37	52.9	317	2	cysteine synthase
23	37	52.9	323	2	conserved hypothet
24	37	52.9	365	2	MHC class I histoc
25	37	52.9	393	2	3-deoxy-manno-octu
26	37	52.9	428	2	histidyl-tRNA synt
27	37	52.9	447	2	hypothetical prote
28	37	52.9	664	2	probable potassium
29	37	52.9	749	2	outer capsid prote

#### ALIGNMENTS

##### RESULT 1

TFHUL

Lactotransferrin precursor [validated] - human

N;Alternate names: lactoferrin

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C;Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61189; A31000; S74

R;Cho, Y.

submitted to the EMBL Data Library, March 1994

A;Reference number: G06820

A;Accession: G01394

A;Status: preliminary; translated from GB/EMBL/DBD

A;Molecule type: mRNA

A;Residues: 1-711 <CHO>

A;Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R;Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A;Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A;Reference number: S11228; MUID:90384839; PMID:2402455

A;Accession: S11228

A;Molecule type: mRNA

A;Residues: 1-148,'T',150-422,'C',424-711 <REY>

A;Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R;Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A;Title: Differential molecular mechanism of the estrogen action that regulates lactofer.

A;Reference number: A45401; MUID:93125571; PMID:1480183

A;Accession: A45401

A;Molecule type: DNA

A;Residues: 1-15 <TEN>

A;Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A;Experimental source: Placenta

A;Note: sequence extracted from NCBI backbone (NCBIP:122202)

R;Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A;Title: Nucleotide sequence of human lactoferrin cDNA.

A;Reference number: S10324; MUID:90326549; PMID:2374734

A;Accession: S10324

A;Molecule type: mRNA

A;Residues: 3-711 <POW>

A;Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R;Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A;Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A;Reference number: S15853; MUID:91264786; PMID:2049066

A;Accession: S15853

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 20-31 <ST1>

A;Accession: S20841

A;Molecule type: protein

A;Residues: 20-28,'X',30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A:Reference number: S07160; MUID:88001031; PMID:3477300  
 A:Accession: S07160  
 A:Molecule type: mRNA  
 A:Residues: 436-487, 'A', 489-711 <RAD>  
 A:CROSS-references: EMBL:ML8642; NID:G186815; PIDN:AAA86665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A:Reference number: A61169; MUID:91235214; PMID:1674448  
 A:Accession: A61169  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 3-701, 'SWKPVN' <PAN>  
 A:Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A>Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A:Reference number: A31000; MUID:85076667; PMID:6510420  
 A:Accession: A31000  
 A:Molecule type: protein  
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A>Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Nørskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity  
 A:Reference number: S74119; MUID:97054624; PMID:8898921  
 A:Accession: S74119  
 A:Molecule type: protein  
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A:Experimental source: neutrophil granulocytes  
 C:Genetics:  
 A:Gene: GDB:ITF  
 A:CROSS-references: GDB:119368; OMIM:150210  
 A:Map position: 3q21-3q23  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <STG>  
 F:20-71/Product: lactotransferrin #status experimental <MAT>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65/39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status e  
 F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:368-400, 378-391, 425-706, 447-669, 479-554, 513-521, 524-537, 847-852/Disulfide bonds: #statu  
 Query Match 97.1%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00049;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQNRNRKV 12  
 DB 39 CFQWQNRNRKV 50  
 RESULT 2  
 lactoferrin - goat  
 C:Species: Capra aegagrus hircus (domestic goat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C:Accession: JC2323  
 R;Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A:Reference number: JC2323; MUID:94380047; PMID:8093048  
 A:Accession: JC2323  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <LEP>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:359-696/Domain: transferrin repeat homology <TRH2>  
 F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 72.9%; Score 51; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.48;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKV 11  
 DB 38 CFQWQNRNRKV 48

## RESULT 3

lactoferrin - sheep (fragment)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C:Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Magliore-Samouri, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A:Reference number: S52107; MUID:95127729; PMID:7827104  
 A:Accession: S52107  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-33 <QIA>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication

Query Match 68.6%; Score 48; DB 2; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.074;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKV 11  
 DB 19 CFQWQNRNRKV 29

## RESULT 4

T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 C:Accession: T08030  
 R;Mitchell, D.R.; Brown, K.S.  
 J. Cell Sci. 107, 635-644, 1994  
 A>Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes  
 A:Reference number: Z16302; MUID:94274778; PMID:8006077  
 A:Accession: T08030  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-4568 <MIT>  
 A:CROSS-references: EMBL:U02963; NID:G409965; PIDN:AAA19956.1; PID:G514215  
 A:Experimental source: strain 21gr  
 C:Genetics:  
 A:Gene: ODA4  
 A:Map position: IX  
 A:Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;  
 3334/3; 3686/3; 3882/3; 4240/3  
 C:Superfamily: dynein heavy chain, ciliary  
 C:Keywords: nucleotide binding; P-loop  
 F:1919-1926/Region: nucleotide-binding motif A (P-loop)  
 F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
 F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 65.7%; Score 46; DB 2; Length 4568;  
 Best Local Similarity 58.3%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNRNRKV 12  
 DB 1852 CFQWQNRNRKV 1863

## RESULT 5

AD2346  
 hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AD2346  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AD2346  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-298 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA876022.1; PID:gl17133459; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr4323

Query Match 60.0%; Score 42; DB 2; Length 298;  
 Best Local Similarity 77.8%; Pred. No. 7.7;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNLK 10  
 |||||  
 Db 163 FHWQNRK 171

RESULT 6  
 A28438  
 lactoferrin precursor - mouse  
 N:Alternate names: lactotransferrin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A28438; A41205  
 R:Penetecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A:Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory gland  
 A:Reference number: A92596; MUID:87280033; PMID:3611056  
 A:Accession: A28438  
 A:Molecule type: mRNA  
 A:Residues: 3-707 <PEN>  
 A:Cross-references: EMBL:J03298  
 R:Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A:Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A:Reference number: A41205; MUID:92042099; PMID:1939212  
 A:Accession: A41205  
 A:Molecule type: DNA  
 A:Residues: 1-15 <LIU>  
 A:Cross-references: GB:M74778  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-707/Product: lactotransferrin #status predicted <MAT>  
 F:358-695/Domain: transferrin repeat homology <TRH2>  
 F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.0%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNLKRV 11  
 |||||  
 Db 37 CURWQNEMRKV 47

RESULT 7  
 AB0858  
 hypothetical protein STY3070 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AB0858  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AB0858  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-511 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:gl6504016; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY3070

Query Match 58.6%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQNLKRV 12  
 |||||  
 Db 350 CFQWQNLKRV 361

RESULT 8  
 T28820  
 hypothetical protein F07C3.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T28820  
 R:Favell, A.; Gattung, S.  
 submitted to the EMBL Data Library, March 1996  
 A:Description: The sequence of C. elegans cosmid F07C3.  
 A:Reference number: Z20528  
 A:Accession: T28820  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-932 <FAV>  
 A:Cross-references: EMBL:U50308; PIDN:AAC48001.1; GSPDB:GN00023; CESP:F07C3.1  
 A:Experimental source: strain Bristol N2; clone F07C3  
 C:Genetics:  
 A:Gene: CESP:F07C3.1  
 A:Map position: 5  
 A:Introns: 14/2; 55/3; 86/2; 159/3; 220/2; 268/3; 327/1; 363/3; 402/3; 463/2; 559/3; 599/3

Query Match 58.6%; Score 41; DB 2; Length 932;  
 Best Local Similarity 63.6%; Pred. No. 37;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNLKRV 12  
 |||||  
 Db 579 FQWQNLKRV 589

RESULT 9  
 E90094  
 26S proteasome SU B5 [imported] - Guillardia theta nucleomorph  
 C:Species: nucleomorph Guillardia theta  
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 C:Accession: E90094  
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reit, A.; Title: The highly reduced genome of an enslaved algal nucleus.  
 Nature 410, 1091-1096, 2001  
 A:Reference number: A99082; MUID:11323671; PMID:11323671  
 A:Accession: E90094  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-205 <DOU>  
 A:Cross-references: GB:AF165818; NID:gl3794510; PIDN:AAK39885.1; GSPDB:GN00150  
 C:Genetics:

A:Gene: prsB5  
A:Map position: 1  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 57.1%; Score 40; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWORNLL 8  
|||:|:|  
Db 63 CFFWERNL 70

## RESULT 10

B60950  
apolipoprotein B-100 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1993 #sequence\_revision 09-Sep-1994 #text\_change 07-Oct-1994  
C:Accession: B60950  
A:Law, A.; Scott, J.  
J. Lipid Res. 31, 1109-1120, 1990  
A:Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL  
A:Reference number: A60950; PMID:90324804; PMID:2373961  
A:Accession: B60950

A:Molecule type: mRNA

A:Residues: 1-274 <LAW>

A:Note: authors translated the codon ATA for residue 8 as Val

C:Superfamily: apolipoprotein B

C:Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 57.1%; Score 40; DB 2; Length 274;

Best Local Similarity 77.8%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WORNLRKVR 12  
|:|:|:|  
Db 39 WDRNRKFR 47

## RESULT 11

F90580  
hypothetical protein MYPV\_5500 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: F90580  
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A:Reference number: A99512; PMID:21267165; PMID:11333084  
A:Accession: F90580

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-282 <KUR>

A:Cross-references: GB:AL445566; PID:GI4089965; PIDN:CAC13723.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 5500

A:Genetic code: SGC3

Query Match 55.7%; Score 39; DB 2; Length 282;

Best Local Similarity 50.0%; Pred. No. 25;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNRKVR 11  
|:|:|:|  
Db 20 FAWQNNIKKI 29

## RESULT 12

TI9429  
hypothetical protein C24H11.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: TI9429  
R;Lloyd, C.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19123

A:Accession: TI9429

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-681 <WIL>

A:Cross-references: EMBL:Z81475; PIDN:CAB03914.1; GSPDB:GN00021; CESP:C24H11.8

A:Experimental source: clone C24H11

C:Genetics:

A:Gene: CESP:C24H11.8

A:Map position: 3

A:Introns: 18/3; 65/3; 108/1; 138/1; 177/3; 256/2; 303/2; 343/1; 409/3; 484/3; 532/2; 58

Query Match 55.7%; Score 39; DB 2; Length 681;

Best Local Similarity 60.0%; Pred. No. 60;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQNRKVR 12  
|:|:|:|  
Db 192 RWQNRKVR 201

## RESULT 13

T22597

hypothetical protein F53H4.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T22597

R;Dobson, R.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19587

A:Accession: T22597

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-275 <WIL>

A:Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4

A:Experimental source: clone F53H4

C:Genetics:

A:Gene: CESP:F53H4.4

A:Map position: X

A:Introns: 67/1; 153/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 54.3%; Score 38; DB 2; Length 275;

Best Local Similarity 54.5%; Pred. No. 36;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNRKVR 12  
|:|:|:|  
Db 262 FQWKSMRKTR 272

## RESULT 14

T22879

hypothetical protein F57G9.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22879

R;Steward, C.

submitted to the EMBL Data Library, December 1996

A:Reference number: Z19631

A:Accession: T22879

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-357 <WIL>

A:Cross-references: EMBL:Z83231; PIDN:CAB05749.1; GSPDB:GN00020; CESP:F57G9.1

A:Experimental source: clone F57G9

C:Genetics:

A:Gene: CESP:F57G9.1

A:Map position: 2

A;Introns: 155/2; 207/3; 248/3; 290/2

Query Match 54.3%; Score 38; DB 2; Length 357;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 FQWQNRKVR 12

Db 317 WKYRRNLKIR 327

# RESULT 15

T00510

probable cytochrome P450 At2g23220 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001

C;Accession: T00510; A84622

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, November 1997

A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.

A;Reference number: Z14164

A;Accession: T00510

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-515 <ROU>

A;Cross-references: EMBL:AC002391; NID:g2642427; PID:g2642441

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

uss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: A84622

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-515 <STO>

A;Cross-references: GB:AE002093; NID:g2642441; PIDN:AAB87109.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g23220; T20D16.15

A;Map position: 2

A;Introns: 182/3; 310/3

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: heme; iron; metalloprotein

F;312-471/Domain: cytochrome P450 homology <P45>

F;449/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match

54.3%; Score 38; DB 2; Length 515;

Best Local Similarity 83.3%; Pred. No. 68;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQR 6

Db 469 CFQWQR 474

Search completed: February 21, 2003, 07:47:54

Job time : 10.65 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-82  
Perfect score: 70  
Sequence: 1 CFQWQRNLKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	97.1	711	1 TREL_HUMAN	P02788 homo sapien
2	51	72.9	708	1 TREL_CAMDR	Q9tum0 camelus dro
3	51	72.9	708	1 TREL_CAPHI	Q29477 capra hircu
4	46	65.7	4568	1 DYHB_CHLRE	Q39565 chlamydomon
5	43	61.4	695	1 TREL_HORSE	O77811 equus cabal
6	42	60.0	707	1 TREL_MOUSE	P08071 mus musculu
7	39	55.7	292	1 NLA_DROME	Q9x218 drosophila
8	38	54.3	62	1 RL28_THETN	Q8r9ul thermoanaer
9	38	54.3	146	1 RPOB_LIBAF	P41187 liberibacte
10	37	52.9	365	1 1A34_HUMAN	P30453 homo sapien
11	37	52.9	428	1 SYH_CHLMU	Q9pj19 chlamydia m
12	37	52.9	749	1 VP4_ROTGA	Q04316 rotavirus (
13	37	52.9	765	1 Y008_HUMAN	Q15398 homo sapien
14	37	52.9	962	1 YBX7_SCHPO	Q10201 schizosacch
15	36	51.4	60	1 RL28_BACST	Q92a12 listeria m
16	36	51.4	62	1 RL28_LISMO	Q23374 bacillus st
17	36	51.4	62	1 RL28_STAAM	Q99up4 staphylococ
18	36	51.4	246	1 Y495_SVNY3	Q55185 synechocyst
19	36	51.4	275	1 IL2A_BOVIN	P12342 bos taurus
20	36	51.4	275	1 IL2A_SHEEP	P26898 ovis aries
21	36	51.4	430	1 SYH_CHLNP	Q9z7p1 chlamydia p
22	36	51.4	453	1 O83A_DROME	Q9vnb3 drosophila
23	36	51.4	455	1 YKVL_CAEEL	P44893 haemophilus
24	36	51.4	475	1 YBBA_HAFIN	Q9ute7 schizosacch
25	36	51.4	496	1 MSL1_SCHPO	Q39017 arabidopsis
26	36	51.4	728	1 KDGI_ARATH	Q21988 caenorhabdi
27	36	51.4	783	1 YNR2_CAEEL	Q9lt02 arabidopsis
28	36	51.4	1179	1 ATX1_ARATH	Q24325 drosophila
29	36	51.4	1213	1 T2D2_DROME	Q14573 homo sapien
30	36	51.4	2671	1 IP3T_HUMAN	P45068 haemophilus
31	35.5	50.7	425	1 FTSA_HAFIN	P37589 salmonella
32	35	50.0	85	1 YMR3_SALTY	Q58878 methanococ
33	35	50.0	152	1 YB83_METJA	

## RESULT 1

ID	TREL_HUMAN	STANDARD;	PRT:	711 AA.
AC	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;			
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrroxin A;			
DE	Lactoferrroxin B; Lactoferrroxin C].			
GN	LTF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 19:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Choi Y.Y.;			
RL	Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Connely O.M.;			
RN	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RL	"Molecular cloning and sequence analysis of human lactoferrin.";			
RT	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RL	"Human neutrophil lactoferrin coding and 5' flanking region DNA sequences.";			
RT	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RL	"cDNA cloning and sequence analysis of human lactoferrin.";			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8]			

34	35	50.0	214	1	VIF_SIVS4	P12505 simian immu
35	35	50.0	235	1	FL3L_HUMAN	P49771 homo sapien
36	35	50.0	238	1	CCDA_BACHD	Q9kdl8 bacillus ha
37	35	50.0	329	1	CATK_RAT	Q35186 rattus norv
38	35	50.0	360	1	HS8_LACLA	Q02135 lactococcus
39	35	50.0	388	1	SH4_HUMAN	Q13639 homo sapien
40	35	50.0	428	1	SYH_CHLTR	O84547 chlamydia t
41	35	50.0	447	1	FD6C_SPIOL	P48629 spinacia ol
42	35	50.0	502	1	C911_ARATH	Q9f965 arabidopsis
43	35	50.0	663	1	PD1L_HUMAN	Q9ulc6 homo sapien
44	35	50.0	704	1	TREL_PIG	P14632 sus scrofa
45	35	50.0	2594	1	7LES_DROVI	P20806 drosophila

## ALIGNMENTS

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogen J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RX Legrand D., Spik G., Montreuil J., Jolles P.;  
RA "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RX Jolles P.;  
RA "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RX Jolles P.;  
RA "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Forcel B.M.,  
RA Dragan Y., Glacalone J., Fae A., Powell E., Solinsky K.A., Desliva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253->methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RL awamori.";  
RA Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RA Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klinthworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hajtancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION. USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERRONIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERRONIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83202; AAA59511.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; M18642; AAA86665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC PDB; 1LCF; 31-AUG-94.  
CC PDB; 1LCT; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFF; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-97.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.



Query Match 97.1%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00017;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
 |||||:||||  
 Db 39 CFQWQNRKVR 50

## RESULT 2

TRFL\_CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9MZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AJ131674; CAB53387.1; -;  
 DR EMBL; AF165879; AAF82241.1; -;  
 DR HSSP; O77811; 1B1X.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PRO0422; TRANSFERRIN.  
 DR SMART; SM00094; TR FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 F -> S (IN REF. 2).  
 FT CONFLICT 261 261 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LJS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 72.9%; Score 51; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.17;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNRKVR 12  
 |||||:||||  
 Db 38 CAQWQNRKVR 49

## RESULT 3

TRFL\_CAPHI STANDARD; PRT; 708 AA.  
 ID TRFL\_CAPHI Q29477; Q29479;  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]\_TaxID=9925;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Nocard M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine UI2 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.





or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL; J03298; AAA0525.1; --  
 CC EMBL; D88510; BAA13633.1; --  
 CC EMBL; BC006904; AAA06904.1; --  
 CC EMBL; M74778; AAA39427.1; --  
 CC PIR; A28438; A28438.  
 CC HSP; P02788; 1CB6.  
 CC MGP; MGI:96837; 1Cf.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR\_FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 CC Transprot; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW signal.  
 CC SIGNAL 1 19 BY SIMILARITY.  
 CC CHAIN 20 707 LACTOTRANSFERRIN.  
 CC REPEAT 20 357 1.  
 CC REPEAT 358 707 2.  
 CC DISULFID 27 63 BY SIMILARITY.  
 CC DISULFID 37 54 BY SIMILARITY.  
 CC DISULFID 133 216 BY SIMILARITY.  
 CC DISULFID 175 191 BY SIMILARITY.  
 CC DISULFID 188 199 BY SIMILARITY.  
 CC DISULFID 249 263 BY SIMILARITY.  
 CC DISULFID 366 398 BY SIMILARITY.  
 CC DISULFID 376 389 BY SIMILARITY.  
 CC DISULFID 423 702 BY SIMILARITY.  
 CC DISULFID 443 665 BY SIMILARITY.  
 CC DISULFID 475 550 BY SIMILARITY.  
 CC DISULFID 499 593 BY SIMILARITY.  
 CC DISULFID 509 523 BY SIMILARITY.  
 CC DISULFID 520 533 BY SIMILARITY.  
 CC DISULFID 591 605 BY SIMILARITY.  
 CC DISULFID 643 648 BY SIMILARITY.  
 CC METAL 78 78 IRON 1 (BY SIMILARITY).  
 CC METAL 110 110 IRON 1 (BY SIMILARITY).  
 CC METAL 210 210 IRON 1 (BY SIMILARITY).  
 CC METAL 271 271 IRON 1 (BY SIMILARITY).  
 CC METAL 413 413 IRON 2 (BY SIMILARITY).  
 CC METAL 451 451 IRON 2 (BY SIMILARITY).  
 CC METAL 544 544 IRON 2 (BY SIMILARITY).  
 CC METAL 613 613 IRON 2 (BY SIMILARITY).  
 CC BINDING 139 139 ANION (POTENTIAL).  
 CC BINDING 481 481 ANION (POTENTIAL).  
 CC CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 494 494 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CONFLICT 1 2 MR -> IOG (IN REF. 1).  
 CC CONFLICT 25 25 R -> Q (IN REF. 2).  
 CC CONFLICT 82 82 M -> L (IN REF. 2).  
 CC CONFLICT 359 359 S -> T (IN REF. 2).  
 CC CONFLICT 382 382 A -> D (IN REF. 1).  
 CC CONFLICT 449 449 E -> G (IN REF. 2).  
 CC CONFLICT 629 629 L -> V (IN REF. 1).  
 CC SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;

Query Match 60.0%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 6.6;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQNRNRKV 11

Db 37 CURWQNRMRKV 47

RESULT 7

NLA\_DROME  
 ID\_NLA\_DROME STANDARD; PRT; 292 AA.

AC Q9XZL8; Q9V391;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 GN NLA OR CG6072.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC McCormick A.V., Goldberg M.L.;  
 CC "Gene required for elongation of meiosis I spindle in Drosophila  
 CC females";  
 CC Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=Berkeley;  
 CC MEDLINE=20196006; PubMed=10731132;  
 CC Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 CC Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 CC George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 CC Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 CC Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 CC Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 CC Abril J.F., Agbayani A., An H.-J., Andrews-pfankuch C., Baldwin D.,  
 CC Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 CC Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 CC Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 CC Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 CC Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 CC de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 CC Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 CC Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 CC Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 CC Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 CC Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 CC Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 CC Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 CC Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 CC Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 CC Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 CC Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 CC Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 CC Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 CC Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 CC Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 CC Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 CC Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 CC Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 CC Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 CC Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 CC Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 CC Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 CC Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 CC "The genome sequence of Drosophila melanogaster";  
 CC Science 287:2185-2195(2000).  
 CC -1- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 CC -1- SIMILARITY: BELONGS TO THE DSCRI FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF147700; AAD33987.1; --  
 EMBL; AE003712; AAF55285.1; --

DR FlyBase; FBgn0026629; nla.

SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;

```

Query Match          55.7%; Score 39; DB 1; Length 292;
Best Local Similarity 54.5%; Pred. No. 8.9;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNLKVR 12
DB 150 FQWLRGFRRLR 160

RESULT 8
RL28_THETN          STANDARD;          PRT;          62 AA.
ID Q8R3U1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L28.
GN RPB6 OR TTE1495.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE013107; AM24713.1; -.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match          54.3%; Score 38; DB 1; Length 62;
Best Local Similarity 60.0%; Pred. No. 2.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQNLKVR 12
DB 27 RWKPNIRKVR 36

RESULT 9
RPOB_LIBAF          STANDARD;          PRT;          146 AA.
ID RPOB_LIBAF
AC P41187;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE beta chain) (RNA polymerase beta subunit) (Fragment).
GN RPOB.
OS Liberibacter africanus (Liberibacter africanum).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Candidatus Liberibacter.
OX NCBI_TaxID=34020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nelspruit;
RA Planet P., Jegoueix S., Bove J.M., Garnier M.;
RT "Detection and characterization of the African Citrus Greening

```

```

RT Liberibacter by amplification, cloning and sequencing of the rplKJL-
RT zpsC operon.";
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U09675; AAA19557.1; -.
DR InterPro; IPR001572; RNA_pol_B.
DR Pfam; PF00562; RNA_pol_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; PARTIAL.
KW Transference; Transcription; DNA-directed RNA polymerase.
FT NON_TER 146
SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FD8943 CRC64;

Query Match          54.3%; Score 38; DB 1; Length 146;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQNLK 10
DB 10 CVQWQSGARK 19

RESULT 10
LA34_HUMAN          STANDARD;          PRT;          365 AA.
ID LA34_HUMAN
AC P30453; P30454;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HLA class I histocompatibility antigen, AW-34 (A-10) alpha chain
DE precursor.
DE HLA-A OR HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=93056508; PubMed=1431115;
RA Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA Tell R.W., du Toit E.D., Farham P.;
RT "Distinctive HLA-A,B antigens of black populations formed by
RT interallelic conversion.";
RL J. Immunol. 149:3411-3415(1992).
RN [2]
RP SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=93235211; PubMed=8475492;
RA Madrigal J.A., Hildebrand W.H., Belich M.P., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA du Toit E.D., Farham P.;
RT "Structural diversity in the HLA-A10 family of alleles: correlations
RT with serology.";
RL Tissue Antigens 41:72-80(1993).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

```



RL J. Virol. 67:2730-2738 (1993).  
 CC 1- SUBCELLULAR LOCATION: Outer capsid.  
 CC 1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

CC EMBL; M91434; AAA47338.1; -;  
 CC InterPro; IPR000416; Cap\_VP4.  
 CC Pfam; PF00426; VP4; 1.  
 CC  
 CC Coat protein; Glycoprotein.  
 CC CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 681 681 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 698 698 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 749 AA; 84362 MW; D1223527DEAE0F21 CRC64;

Query Match 52.9%; Score 37; DB 1; Length 749;  
 Best Local Similarity 50.0%; Pred. No. 53;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CFQWQNLKRV 12  
 ||:|:|:|  
 DB 195 CFWDNCAVNR 206

RESULT 13  
 Y008 HUMAN  
 ID Y008 HUMAN STANDARD; PRT; 765 AA.  
 AC Q15398;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein KIAA0008.  
 GN KIAA0008  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC  
 CC NCB1\_TaxID=9606;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE=Bone marrow;  
 CC MEDLINE=96051387; PubMed=7584026;  
 CC Nomura N., Miyajima N., Suzuki T., Tanaka A., Kawarabayashi Y.,  
 CC Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;  
 CC "Prediction of the coding sequences of unidentified human genes. I.  
 CC The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by  
 CC analysis of randomly sampled cDNA clones from human immature myeloid  
 CC cell line KG-1.";  
 CC DNA Res. 1:27-35 (1994).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE=Eye, and Lung;  
 CC Strausberg R.;  
 CC Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC

CC EMBL; D13633; BAA02977.1; -;  
 CC EMBL; BC010658; AAH10658.1; -;  
 CC EMBL; BC016276; AAH16276.1; -;  
 CC InterPro; IPR005026; GKAP.  
 CC Pfam; PF03359; GKAP; 1.  
 CC KW Hypothetical protein.  
 CC SQ SEQUENCE 765 AA; 85668 MW; 00AFF91A02387EAL CRC64;

Query Match 52.9%; Score 37; DB 1; Length 765;  
 Best Local Similarity 62.5%; Pred. No. 54;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNL 8  
 ||:|:|:|  
 DB 367 CFWDRL 374

RESULT 14  
 YBX7 SCHPO  
 ID YBX7 SCHPO STANDARD; PRT; 962 AA.  
 AC Q10201;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C17D1.07c in chromosome II.  
 GN SPB17D1.07C.

OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 CC NCB1\_TaxID=4896;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=972;  
 CC MEDLINE=21848401; PubMed=11859360;

CC Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 CC Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 CC Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 CC Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 CC Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 CC Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 CC James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 CC Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 CC Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 CC Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 CC Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 CC Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 CC Woodward J., Voicakert G., Aert R., Robben J., Grymonprez B.,  
 CC Weltjens I., Vanstreels E., Kieger M., Schaefer M., Mueller-Auer S.,  
 CC Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 CC Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 CC Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 CC Goffeau A., Cadieu B., Dreano S., Gloux C., Lelaure V., Mottier S.,  
 CC Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 CC Lucas M., Rochet M., Gailardin C., Tallada V.A., Garzon A., Thode G.,  
 CC Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 CC Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 CC Cerrutti I., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 CC Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 CC "The genome sequence of Schizosaccharomyces pombe.";  
 CC Nature 415:871-880 (2002).  
 CC -!- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; AL031322; CAA20431.1; --  
 DR InterPro; IPR001715; Calponin-like.  
 DR InterPro; IPR000048; IQ\_region.  
 DR Pfam; PF00612; IQ; 1.  
 DR ProDom; PD001527; CH\_type; 1.  
 DR SMART; SM00033; CH; 1.  
 DR PROSITE; PS00021; CH; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 64 170 CH.  
 SQ SEQUENCE 962 AA; 112678 MW; 97FD7FD9942DEB CRC64;

Query Match 52.9%; Score 37; DB 1; Length 962;  
 Best Local Similarity 66.7%; Pred. No. 69;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQRLKRV 11  
 DB 756 QWQSLRKI 764

## RESULT 15

RL28\_BACST STANDARD; PRT; 60 AA.  
 AC P23374;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 50S ribosomal protein L28.  
 GN RPMB.  
 OS Bacillus stearothermophilus.  
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.  
 OX NCBI\_TaxID=1422;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92075758; PubMed=1742360;  
 RA Kruff V., Kapp U., Wittmann-Liebold B.;  
 RT "Characterization and primary structure of proteins L28, L33 and L34  
 from Bacillus stearothermophilus ribosomes."  
 RL Biochimie 73:855-860(1991).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 DR PIR; A48396; A48396.  
 DR InterPro; IPR001383; Ribosomal\_L28.  
 DR Pfam; PF00830; Ribosomal\_L28; 1.  
 DR TIGRFAMs; TIGR00009; L28; 1.  
 KW Ribosomal protein.  
 FT INIT MET 0  
 SQ SEQUENCE 60 AA; 6810 MW; 2AD9161CD60E92F4 CRC64;

Query Match 51.4%; Score 36; DB 1; Length 60;  
 Best Local Similarity 66.7%; Pred. No. 5.8;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QWQRLKRV 12  
 DB 27 WKANLQKVR 35

Search completed: February 21, 2003, 07:27:57  
 Job time : 4.6 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107B-82

Perfect score: 70 CFQWQNRNLRKV 12

Sequence: 1 CFQWQNRNLRKV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	90.0	711	4 Q8TCD2	Q8Tcd2 homo sapien
2	59	84.3	38	4 Q9UCY5	Q9ucy5 homo sapien
3	48	68.6	33	6 Q9TR80	Q9tr80 ovine aries
4	42	60.0	298	16 Q8YP77	Q8yp77 anabaena sp
5	42	60.0	469	9 Q38115	Q38115 bacterioph
6	41	58.6	273	2 Q31090	Q31090 rhizobium l
7	41	58.6	306	4 Q8TAX2	Q8tax2 homo sapien
8	41	58.6	466	4 Q8NUS2	Q8nus2 homo sapien
9	41	58.6	488	10 Q8S934	Q8s934 diospyros k
10	41	58.6	511	16 Q8Z462	Q8z462 salmonella
11	41	58.6	932	5 Q19153	Q19153 caenorhabdi
12	40	57.1	91	15 Q77855	Q77855 human immun
13	40	57.1	91	15 Q77856	Q77856 human immun
14	40	57.1	205	8 Q8RRR2	Q8rrr2 guillardia
15	40	57.1	279	16 Q8XS52	Q8xs52 raietonia s
16	40	57.1	570	10 Q8S487	Q8s487 zea mays (m

#### ALIGNMENTS

##### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; --  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 90.0%; Score 63; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0046;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNRNLRKV 11  
Db 39 CFQWQNRNLRKV 49

##### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96081613; PubMed=8551695;  
RA Sato I.;  
RT "Characterization of the 84-kDa protein with ABH activity in human  
RL seminal plasma.";  
RL Jpn. J. Legal Med. 49:281-293 (1995).  
DR HSPF; P02788; 1BKA.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDEB CRC64;  
  
Query Match 84.3%; Score 59; DB 4; Length 38;  
Best Local Similarity 90.9%; Pred. No. 0.0011;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 FQWQNRKVR 12  
DB 21 FQWQNRKVR 31  
  
RESULT 3  
Q9TR80 PRELIMINARY; PRT; 33 AA.  
ID Q9TR80  
AC Q9TR80;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95127729; PubMed=7827104;  
RA Qian Z.-Y., Colles P., Migliore-Samour D., Fiat A.M.;  
RL Biochim. Biophys. Acta 1243:25-32 (1995).  
DR HSPF; O77698; 1CE2.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;  
  
Query Match 68.6%; Score 48; DB 6; Length 33;  
Best Local Similarity 54.5%; Pred. No. 0.09;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CFQWQNRKVR 11  
DB 19 CFQWQNRKVR 29  
  
RESULT 4  
Q8YP77 PRELIMINARY; PRT; 298 AA.  
ID Q8YP77  
AC Q8YP77;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein Alr4323.  
GN ALR4323.  
OC Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yanada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium Anabaena sp. strain PCC 7120.";  
RL DNA Res. 8:205-213 (2001).  
DR EMBL; AP003596; BAB76022.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 298 AA; 34513 MW; 64036E6B52299A9F CRC64;  
  
Query Match 60.0%; Score 42; DB 16; Length 298;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 FQWQNRKVR 10  
DB 163 FQWQNRKVR 171  
  
RESULT 5  
Q38115 PRELIMINARY; PRT; 469 AA.  
ID Q38115  
AC Q38115;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE ORF29.  
OS Bacteriophage rlt.  
OC Viruses.  
OX NCBI\_TaxID=43685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96332669; PubMed=8730874;  
RA Nauta A.; Van Sinderen D.; Karsens H., Smit E., Venema G., Kok J.;  
RT "Inducible gene expression mediated by a repressor-operator system  
RT isolated from Lactococcus lactis bacteriophage rlt.";  
RL Mol. Microbiol. 19:1331-1341 (1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96332669; PubMed=8730875;  
RA Van Sinderen D., Karsens H., Kok J., Terpstra P., Ruiters M.H.,  
RA Venema G., Nauta A.;  
RT "Sequence analysis and molecular characterization of the temperate  
RT lactococcal bacteriophage rlt.";  
RL Mol. Microbiol. 19:1343-1355 (1996).  
DR EMBL; U89906; AAB18704.1; -.  
SQ SEQUENCE 469 AA; 53160 MW; 1F8E02D4325C6BB9 CRC64;  
  
Query Match 60.0%; Score 42; DB 9; Length 469;  
Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 CFQWQNRKVR 10  
DB 39 CYPWQNRKVR 48  
  
RESULT 6  
O31090 PRELIMINARY; PRT; 273 AA.  
ID O31090  
AC O31090;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE Hypothetical 31.0 kDa protein.  
OS Rhizobium leguminosarum (biovar viciae).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF39;  
RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;  
RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF39;  
 RX MEDLINE=99113394; PubMed=9914965;  
 RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,  
 RA Ksenenko V.N.;  
 RT "Structural and functional organization of the exopolysaccharide  
 RT biosynthesis genes in *Rhizobium leguminosarum* bv. *viciae* VF39.";  
 RL Mol. Biol. (Mosk) 32:797-804(1998).  
 DR EMBL; AF028810; AAB88891.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 273 AA; 30986 MW; F195D2FEF7AD44D3 CRC64;

Query Match 58.6%; Score 41; DB 2; Length 273;  
 Best Local Similarity 70.0%; Pred. No. 16;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQWQNLKVR 12  
 Db 245 RWLNLKLR 254

RESULT 7  
 Q8TAX2 PRELIMINARY; PRT; 306 AA.

ID Q8TAX2  
 AC Q8TAX2;  
 DT 01-JUN-2002 (TREMELrel. 21, Created)  
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE Similar to hypothetical protein FJ11175.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANCREAS;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025708; AAH25708.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 306 AA; 43D9AD722566B343 CRC64;

Query Match 58.6%; Score 41; DB 4; Length 306;  
 Best Local Similarity 66.7%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNLKVR 9  
 Db 269 CFQWQNLKVR 277

RESULT 8  
 Q9NUS2 PRELIMINARY; PRT; 466 AA.

ID Q9NUS2  
 AC Q9NUS2;  
 DT 01-OCT-2000 (TREMELrel. 15, Created)  
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE CDNA FJ11175.1; clone pLACE1007375, weakly similar to phorbol  
 DE ester/diacylglycerol-binding protein UNC-13.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK002037; BAA92048.1; -  
 DR HSSP; P21707; 1BYN.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00168; C2; 1  
 DR PRINTS; PR00360; C2DOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
 DR PROSITE; PS00499; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 SQ SEQUENCE 466 AA; 53192 MW; E4113A5062F58D6E CRC64;

Query Match 58.6%; Score 41; DB 4; Length 466;  
 Best Local Similarity 66.7%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNLKVR 9  
 Db 269 CFQWQNLKVR 277

RESULT 9

ID Q8S934 PRELIMINARY; PRT; 488 AA.  
 AC Q8S934;  
 DT 01-JUN-2002 (TREMELrel. 21, Created)  
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE 1-aminocyclopropane-1-carboxylate synthase.  
 GN DK-ACS2.  
 OS Diospyros kaki (kaki persimmon).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Ericales; Ebenaceae; Diospyros.  
 OX NCBI\_TaxID=35925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. HIRATANENASHI;  
 RA Nakano R., Ogura E., Kubo Y., Inaba A.;  
 RT "Water stress induces ethylene biosynthesis in Japanese persimmon  
 RT fruit.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB073006; BAB89349.1; -  
 SQ SEQUENCE 488 AA; 55238 MW; 81C38BE8F67C21AD CRC64;

Query Match 58.6%; Score 41; DB 10; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QWQWQNLKVR 9  
 Db 452 QWQWQNLKVR 458

RESULT 10

ID Q8Z462 PRELIMINARY; PRT; 511 AA.  
 AC Q8Z462;  
 DT 01-MAR-2002 (TREMELrel. 20, Created)  
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)  
 DE Hypothetical protein STY3070.  
 GN STY3070.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).;
DR EMBL; AL627276; CAD6049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DDDI2410D178B CRC64;

Query Match 58.6%; Score 41; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFQWQNLKRV 12
Db 350 CFANDMKAKV 361
|||:|||||

RESULT 11
Q19153 PRELIMINARY; PRT; 932 AA.
AC Q19153;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 105.1 kDa protein.
F07C3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=95069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Pavello A., Gattung S.;
RT "The sequence of C. elegans cosmid F07C3.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50308; AAG24025.1; -.
DR InterPro; IPR000731; HMGCR/patch_5TM.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 932 AA; 105144 MW; 66680619ADACBFD5 CRC64;

Query Match 58.6%; Score 41; DB 5; Length 932;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FQWQNLKRV 12
Db 579 FQWQSRARLVK 589
|||:|||||

STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).;
DR EMBL; AL627276; CAD6049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DDDI2410D178B CRC64;

Query Match 58.6%; Score 41; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFQWQNLKRV 12
Db 350 CFANDMKAKV 361
|||:|||||

RESULT 12
Q77855 PRELIMINARY; PRT; 91 AA.
AC Q77855;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95191002; PubMed=7884875;
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
RA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
RT "Similarity in env and gag genes between genomic RNAs of human
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is
RT unrelated to time of HIV-1 RNA positivity in the child.";
RL J. Virol. 69:2285-2296(1995).
DR EMBL; Z47867; CAA87881.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 91
SQ SEQUENCE 91 AA; 10530 MW; 8B10C62011F305D6 CRC64;

Query Match 57.1%; Score 40; DB 15; Length 91;
Best Local Similarity 70.0%; Pred. No. 7.4;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QWQNLKRV 12
Db 67 QWNTLQKV 76
|||:|||||

RESULT 13
Q77856 PRELIMINARY; PRT; 91 AA.
AC Q77856;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95191002; PubMed=7884875;
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
RA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
RT "Similarity in env and gag genes between genomic RNAs of human
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is
RT unrelated to time of HIV-1 RNA positivity in the child.";
RL J. Virol. 69:2285-2296(1995).
DR EMBL; Z47868; CAA87882.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 91
SQ SEQUENCE 91 AA; 10519 MW; EB20C63A22DA1288 CRC64;

Query Match 57.1%; Score 40; DB 15; Length 91;
Best Local Similarity 70.0%; Pred. No. 7.4;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QWQNLKRV 12

```

Query Match 57.1%; Score 40; DB 16; Length 279;  
 Best Local Similarity 41.7%; Pred. No. 25;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQNLKVR 12  
 |||:|  
 Db 244 CFQWKEGTAK 255  
 |||:|

Search completed: February 21, 2003, 07:44:36  
 Job time : 20.8 secs

Db 67 QWDRTLQKVR 76  
 |||:|

RESULT 14

Q98RR2 PRELIMINARY; PRT; 205 AA.  
 AC Q98RR2;  
 DT 01-OCT-2001 (TRENBLrel. 18, Created)  
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE 26S proteasome SU B5.  
 GN PRS85.  
 OS Guillardia theta (Cryptomonas phi).  
 OG Nucleomorph.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 CX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21223349; PubMed=11323671;  
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,  
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;  
 RT "The highly reduced genome of an enslaved algal nucleus."  
 RL Nature 410:1091-1096(2001).  
 DR EMBL; AF165818; AAK39885.1; --  
 DR InterPro; IPR001353; Proteasome B.  
 DR Pfam; PF00227; Proteasome; 1.  
 DR PRINTS; PR00141; PROTEASOME.  
 KW Proteasome.  
 SQ SEQUENCE 205 AA; 22691 MW; D30F5289CBC85049 CRC64;

Query Match 57.1%; Score 40; DB 8; Length 205;  
 Best Local Similarity 75.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNL 8  
 |||:|  
 Db 63 CFFWERNL 70  
 |||:|

RESULT 15

Q8XSE2 PRELIMINARY; PRT; 279 AA.  
 AC Q8XSE2;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Putative ICC protein homolog.  
 GN ICC OR RSP0534 OR RS00414.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OG Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 CX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siquier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646079; CAD17685.1; --  
 DR InterPro; IPR004843; M-pppestrase.  
 DR Pfam; PF00149; Metallophos; 1.  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 279 AA; 31541 MW; ABH38818004B2EDA CRC64;

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107B-83  
Perfect score: 71  
Sequence: 1 CFQWQRMKKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	100.0	12	21	AAV78083 Human lactoferrin
2	68	95.8	12	21	AAV78038 Human lactoferrin
3	68	95.8	12	21	AAV78046 Human lactoferrin
4	68	95.8	12	21	AAV78047 Human lactoferrin
5	68	95.8	12	21	AAV78037 Human lactoferrin
6	68	95.8	13	21	AAV78048 Human lactoferrin
7	68	95.8	13	21	AAV78049 Human lactoferrin
8	68	95.8	14	21	AAV78036 Human lactoferrin
9	68	95.8	14	21	AAV78050 Human lactoferrin
10	68	95.8	14	21	AAV78051 Human lactoferrin

11	68	95.8	15	17	AAV98554 Peptide for anti-u
12	68	95.8	15	21	AAV78035 Human lactoferrin
13	68	95.8	15	21	AAV78062 Human lactoferrin
14	68	95.8	15	21	AAV78063 Human lactoferrin
15	68	95.8	16	21	AAV78031 Human lactoferrin
16	68	95.8	16	21	AAV78064 Human lactoferrin
17	68	95.8	16	21	AAV78065 Human lactoferrin
18	68	95.8	17	21	AAV78034 Human lactoferrin
19	68	95.8	17	21	AAV78066 Human lactoferrin
20	68	95.8	17	21	AAV78067 Human lactoferrin
21	68	95.8	18	15	AAV69352 Human lactoferrin
22	68	95.8	18	17	AAV13397 Human lactoferrin
23	68	95.8	18	21	AAV78033 Human lactoferrin
24	68	95.8	19	21	AAV68867 Amino acid sequenc
25	68	95.8	19	21	AAV78032 Human lactoferrin
26	68	95.8	20	13	AAV21810 Anti microbial pep
27	68	95.8	20	14	AAV44841 Lactoferrin-relate
28	68	95.8	20	15	AAV48530 Lactoferrin derive
29	68	95.8	20	15	AAV48531 Lactoferrin derive
30	68	95.8	20	15	AAV57461 Lactoferrin derive
31	68	95.8	20	15	AAV57462 Lactoferrin derive
32	68	95.8	20	16	AAV84698 Bovine lactoferrin
33	68	95.8	20	16	AAV84699 Bovine lactoferrin
34	68	95.8	20	16	AAV80263 Anti-parasitic lac
35	68	95.8	20	16	AAV80264 Anti-parasitic lac
36	68	95.8	20	17	AAV98553 Peptide for anti-u
37	68	95.8	20	17	AAV91852 Lactoferrin-derive
38	68	95.8	20	17	AAV03045 Lactoferrin-derive
39	68	95.8	20	17	AAV90607 Lactoferrin-derive
40	68	95.8	20	17	AAV87621 Lactoferrin-derive
41	68	95.8	20	17	AAV26150 Lactoferrin-derive
42	68	95.8	20	18	AAV26150 Lactoferrin-derive
43	68	95.8	20	18	AAV14036 Anti-parasitic pep
44	68	95.8	20	19	AAV70310 Thrombus formation
45	68	95.8	20	19	AAV53224 Lactoferrin hydrol

## ALIGNMENTS

RESULT 1  
AAV78083  
ID AAV78083 standard; Peptide; 12 AA.  
XX  
AC AAV78083;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:83.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
Query Match 100.0%; Score 71; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 8.3e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQRMKKVR 12  
Db 1 CFOWQRMKKVR 12  
|||||:|||||

RESULT 2  
AAY78038  
ID AAY78038 standard; Peptide; 12 AA.  
XX AAY78038;  
AC AAY78038;  
DT 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:38.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1999; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 12; Page 70; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
Query Match 95.8%; Score 68; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 2.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQRMKKVR 12  
Db 1 CFOWQRMKKVR 12  
|||||:|||||

RESULT 3  
AAY78046  
ID AAY78046 standard; Peptide; 12 AA.  
XX AAY78046;  
AC AAY78046;  
DT 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:46.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1999; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 15; Page 35; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 2.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRNNMKVR 12  
 |||||:|  
 Db 1 CFQWQRNNMKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:47.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX FN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 2.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRNNMKVR 12  
 |||||:|  
 Db 1 CFQWQRNNMKVR 12

RESULT 5  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:37.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX FN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 3e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 1 CFQWRNMKKVR 12  
 |||||:|  
 Db 2 CFQWRNMKKVR 13

## RESULT 6

AAV78048  
 ID AAV78048 standard; Peptide; 13 AA.

AC AAV78048;  
 XX  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 74; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 3e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMKKVR 12  
 |||||:|  
 Db 2 CFQWRNMKKVR 13

## RESULT 7

AAV78049

ID AAV78049 standard; Peptide; 13 AA.  
 XX  
 AC AAV78049;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 3e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMKKVR 12  
 |||||:|  
 Db 2 CFQWRNMKKVR 13

## RESULT 8

AAV78036

ID AAV78036 standard; Peptide; 14 AA.  
 XX  
 AC AAV78036;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCII-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 69; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 14 AA;  
 SQ Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 3.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRNMKKVR 12  
 DB 3 CFQWQRNMKKVR 14  
 RESULT 9  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AAY78050;  
 AC AAY78051;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:50.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.

OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCII-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 75; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 14 AA;  
 SQ Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 3.2e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWQRNMKKVR 12  
 DB 3 CFQWQRNMKKVR 14  
 RESULT 10  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX AAY78051;  
 AC AAY78051;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:51.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF

```

XX 06-JUL-1998;    98SE-0002441.
XX 17-JUL-1998;    98SE-0002562.
XX 29-DEC-1998;    98SE-0004614.
XX
XX (ASCII-) A+ SCI INVEST AB.
XX
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX
XX WPI; 2000-147388/13.
XX
XX New peptides used for treatment and prevention of infections, -
XX inflammations and tumors and for use in infant formula food
XX
XX Claim 18; Page 75; 102pp; English.
XX
XX AAY78001 to AAY78100 represent peptides having sequences based on human
XX lactoferrin. The peptides are taken up in the intestine through
XX binding to specific lactoferrin receptors and are then transported
XX through the circulation. A medicinal product of the peptide or fragment
XX can be used for treating and/or prevention of infections (such as
XX urinary tract infections, colitis, and Candida infection on a mucosal
XX membrane), inflammations and/or tumours. The peptides can also be used
XX in food stuffs such as infant formula food. The peptides are also
XX fungicidal and bactericidal and may also be used as preservatives.
XX Even though native human lactoferrin have been shown to have desired
XX anti-inflammatory anti-infectious and anti-tumoural properties they
XX cannot be used clinically on a broad basis because of high production
XX costs. Therefore, provision of peptides based on lactoferrin would
XX enable them to be used for the same purposes as lactoferrin at lower
XX cost.
XX
XX SQ Sequence 14 AA;
XX
XX Query Match          95.8%; Score 68; DB 21; Length 14;
XX Best Local Similarity 91.7%; Pred.No. 3.2e-05;
XX Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY      1 CFQWQRNKKVR 12
XX         |||||::|||
XX         |||||::|||
XX Db       3 CFQWQRNKKVR 14
XX
XX RESULT 11
XX AAR98554
XX ID AAR98554 standard; Peptide; 15 AA.
XX AC AAR98554;
XX XX
XX DT 12-NOV-1996 (first entry)
XX
XX DE Peptide for anti-ulcer agent.
XX
XX KW anti-ulcer agent; low toxicity; stable; heat-resistant.
XX
XX OS Synthetic.
XX
XX XX JP08143468-A.
XX PN
XX XX 04-JUN-1996.
XX PD
XX PF 17-NOV-1994; 94JP-0283869.
XX PP
XX XX 17-NOV-1994; 94JP-0283869.
XX PR
XX XX (MORG ) MORINAGA MILK IND CO LTD.
XX PA
XX XX WPI; 1996-318857/32.
XX DR
XX XX
XX XX Anti-ulcer agent contg. peptide - has low toxicity, is
XX PT heat-resistant and water-soluble
XX PS Claim 1; Page 11; libpp; Japanese.

```

XX	AAR9531-54	are peptides used in an anti-ulcer agent. The agent is low
CC	in toxicity,	is heat-resistant and stable in aqueous soln.. It can be
CC	administered orally	and be produced in large amounts.
XX		
XX	Sequence	15 AA;
XX		
XX	Query Match	95.8%; Score 68; DB 17; Length 15;
XX	Best Local Similarity	91.7%; Pred. No. 3.5e-05;
XX	Matches	11; Conservative 1; Mismatches 0; Indels 0; Gaps
XX		
XX	QY	1 CFQWRNMKKVR 12
XX		:
XX	Db	2 CFQWRNMKKVR 13
XX		
XX	RESULT 12	
XX	AAV78035	
XX	ID	AAV78035 standard; Peptide; 15 AA.
XX	AC	AAV78035;
XX	XX	
XX	DT	25-APR-2000 (first entry)
XX	XX	
XX	DE	Human lactoferrin derived peptide SEQ ID NO:35.
XX	XX	
XX	KW	Human; lactoferrin; modification; infection; inflammation; tumour;
XX	KW	food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX	KW	urinary tract infection; colitis; Candida infection; fungicidal;
XX	KW	bactericidal; preservative.
XX	XX	
XX	OS	Homo sapiens.
XX	OS	Synthetic.
XX	XX	
XX	PN	WO200001730-A1.
XX	XX	
XX	PD	13-JAN-2000.
XX	XX	
XX	PF	06-JUL-1999; 99WO-SE01230.
XX	XX	
XX	PR	06-JUL-1998; 98SE-0002441.
XX	PR	17-JUL-1998; 98SE-0002562.
XX	PR	29-DEC-1998; 98SE-0004614.
XX	XX	
XX	PA	(ASCII-) A+ SCI INVEST AB.
XX	XX	
XX	PI	Hanson IA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX	XX	
XX	DR	WPI; 2000-147388/13.
XX	XX	
XX	PT	New peptides used for treatment and prevention of infections,
XX	PT	inflammations and tumors and for use in infant formula food -
XX	XX	
XX	PS	Claim 12; Page 69; 102pp; English.
XX	XX	
XX	CC	AAV78001 to AAV78100 represent peptides having sequences based on human
XX	CC	lactoferrin. The peptides are taken up in the intestine through
XX	CC	binding to specific lactoferrin receptors and are then transported
XX	CC	through the circulation. A medicinal product of the peptide or fragment
XX	CC	can be used for treating and/or prevention of infections (such as
XX	CC	urinary tract infections, colitis, and Candida infection on a mucosal
XX	CC	membrane), inflammations and/or tumours. The peptides can also be used
XX	CC	in food stuffs such as infant formula food. The peptides are also
XX	CC	fungicidal and bactericidal and may also be used as preservatives.
XX	CC	Even though native human lactoferrin have been shown to have desired
XX	CC	anti-inflammatory anti-infectious and anti-tumoural properties they
XX	CC	cannot be used clinically on a broad basis because of high production
XX	CC	costs. Therefore, provision of peptides based on lactoferrin would
XX	CC	enable them to be used for the same purposes as lactoferrin at lower
XX	CC	cost.
XX	XX	
XX	XX	Sequence 15 AA;
XX	SQ	

XX AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
CC administered orally and be produced in large amounts.  
CC

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 3.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12  
 DB 4 CFQWQRMKKVR 15

## RESULT 13

AAV78062  
 ID AAV78062 standard; Peptide; 15 AA.

XX AC AAV78062;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:62.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 81; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 3.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12  
 DB 4 CFQWQRMKKVR 15

## RESULT 14

AAV78063  
 ID AAV78063 standard; Peptide; 15 AA.

XX AC AAV78063;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:63.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 81; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 3.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12  
 DB 4 CFQWQRMKKVR 15

## RESULT 15

AAV78031  
 ID AAV78031 standard; Peptide; 16 AA.

XX AC AAV78031;

XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:31.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PW 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR New peptides used for treatment and prevention of infections,  
XX inflammations and tumors and for use in infant formula food -  
PT Claim 11; Page 68; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX SQ Sequence 16 AA;  
Query Match 95.8%; Score 68; DB 21; Length 16;  
Best Local Similarity 91.7%; Pred. NO. 3.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWQRNKKVR 12  
DB 5 CFQWQRNKKVR 16

Search completed: February 21, 2003, 07:37:14  
Job time : 28.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-83  
Perfect score: 71  
Sequence: 1 CFQWQRNMKKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	18	1	US-08-204-487-3
2	68	95.8	18	2	US-08-485-948-8
3	68	95.8	18	2	US-08-628-380-8
4	68	95.8	18	2	US-08-475-055-8
5	68	95.8	20	1	US-07-755-161A-3
6	68	95.8	20	1	US-07-891-174-3
7	68	95.8	20	1	US-08-204-487-1
8	68	95.8	20	1	US-08-256-771-24
9	68	95.8	20	1	US-08-256-771-25
10	68	95.8	20	1	US-08-381-984-24
11	68	95.8	20	1	US-08-381-984-25
12	68	95.8	22	4	US-09-508-734-4
13	68	95.8	24	4	US-09-508-734-6
14	68	95.8	25	1	US-07-755-161A-10
15	68	95.8	25	1	US-07-891-174-10
16	68	95.8	25	1	US-08-204-487-7
17	68	95.8	29	4	US-09-508-734-8
18	68	95.8	36	1	US-07-755-161A-8
19	68	95.8	36	1	US-07-891-174-8
20	68	95.8	36	1	US-08-256-771-30
21	68	95.8	36	1	US-08-381-984-29
22	68	95.8	47	2	US-08-464-182A-6
23	68	95.8	47	2	US-08-406-271-6
24	68	95.8	50	2	US-08-693-274A-7
25	68	95.8	52	4	US-09-017-043A-3
26	68	95.8	53	2	US-08-464-182A-5
27	68	95.8	53	2	US-08-406-271-5

28	68	95.8	54	2	US-08-464-182A-2
29	68	95.8	54	2	US-08-406-271-2
30	68	95.8	694	3	US-08-724-586-2
31	68	95.8	694	4	US-09-421-632-2
32	68	95.8	694	4	US-09-932-190-2
33	68	95.8	705	2	US-08-655-640-2
34	68	95.8	708	2	US-08-655-640-4
35	68	95.8	711	1	US-08-154-019-4
36	68	95.8	711	1	US-08-461-333-4
37	68	95.8	711	3	US-08-464-167-4
38	68	95.8	711	3	US-09-158-313-4
39	68	95.8	711	4	US-08-476-798-4
40	65	91.5	711	1	US-08-145-681-2
41	65	91.5	711	1	US-08-250-308-2
42	65	91.5	711	1	US-08-453-703-2
43	65	91.5	711	2	US-08-456-106-2
44	65	91.5	711	3	US-08-456-108-2
45	65	91.5	711	4	US-09-265-577-2

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5585425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERAKI  
; APPLICANT: DOSAKO, SHUN-ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HORWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 95.8%; Score 68; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 2.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNMKKVR 12  
Db 1 CFQWQNMKKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 2.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNMKKVR 12  
Db 1 CFQWQNMKKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 2.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQNMKKVR 12  
Db 1 CFQWQNMKKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;;  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-C1, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;;  
US-08-475-055-8

Query Match 95.8%; Score 68; DB 2; Length 10;  
Best Local Similarity 91.74; Pred. No. 2.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNKKVR 12  
Db 1 CFQWRNKKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
;; US-07-755-161A-3

Query Match 95.8%; Score 68; DB 1; Length 20;



```
Best Local Similarity 91.7%; Pred. No. 2.5e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNKKVR 12
   |||||:||||
DB 2 CFQWQNNKKVR 13

RESULT 6
US-07-891-174-3
; Sequence 3, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site

; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-3

Query Match 95.8%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 2.5e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNNKKVR 12
   |||||:||||
DB 2 CFQWQNNKKVR 13

RESULT 7
US-08-204-487-1
; Sequence 1, Application US/08204487
; Patent No. 5565425
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, NAOKI
; APPLICANT: NAKASHIMA, HIDEKI
; APPLICANT: MOSUCHI, WATARU
; APPLICANT: TANAKA, SHIGEKI
; APPLICANT: DOSAKO, SHUN'ICHI
; APPLICANT: KAWASAKI, YOSHIHIRO
; APPLICANT: UCHIDA, TOSHIAKI
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,487
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FUN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
```

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWRNMKKVR 12  
Db 2 CFQWRNMKKVR 13

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMIYA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"  
OTHER INFORMATION:

US-08-256-771-24

Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWRNMKKVR 12  
Db 2 CFQWRNMKKVR 13

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMIYA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
OTHER INFORMATION:  
US-08-256-771-25

Query Match 95.8%; Score 68; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWRNMKKVR 12  
Db 2 CFQWRNMKKVR 13

RESULT 10  
US-08-381-984-24

; Sequence 24, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
; US-08-381-984-24  
; Query Match 95.8%; Score 68; DB 1; Length 20;  
; Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
; QY 1 CFQWQRNMKKVR 12  
; Db 2 CFQWQRNMKKVR 13  
; RESULT 11  
; US-08-381-984-25  
; Sequence 25, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack

; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"  
; US-08-381-984-25  
; Query Match 95.8%; Score 68; DB 1; Length 20;  
; Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
; QY 1 CFQWQRNMKKVR 12  
; Db 2 CFQWQRNMKKVR 13  
; RESULT 12  
; US-09-508-734-4  
; Sequence 4, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; TITLE OF INVENTION: useful microorganism thereof  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KR99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: KopatentIn 1.71  
; SEQ ID NO 4

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match          95.8%; Score 68; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 2.7e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMKKVR 12
   |||||:||||
Db 2 CFQWQRNMKKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          95.8%; Score 68; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 3e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMKKVR 12
   |||||:||||
Db 3 CFQWQRNMKKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.,
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLER:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-755-161A-10

Query Match          95.8%; Score 68; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 3.1e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMKKVR 12
   |||||:||||
Db 4 CFQWQRNMKKVR 15
```

## RESULT 15

```
US-07-891-174-10
; Sequence 10, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLAE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21"
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
```

```
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
; Query Match 95.8%; Score 68; DB 1; Length 25;
; Best Local Similarity 91.7%; Pred. No. 3.1e-05;
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
; QY 1 CFQWQRMKKVR 12
; Db 4 CFQWQRMKKVR 15
; Search completed: February 21, 2003, 07:50:36
; Job time : 8.7 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-83  
Perfect score: 71  
Sequence: 1 CFQQRNMKKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	68	95.8	15	9	US-09-798-869-2
2	68	95.8	25	9	US-09-798-869-20
3	68	95.8	694	9	US-10-023-096-2
4	60	84.5	15	9	US-09-798-869-6
5	51	71.8	15	9	US-09-798-869-3
6	51	71.8	25	9	US-09-798-869-23
7	45	63.4	15	9	US-09-798-869-29
8	45	63.4	15	9	US-09-798-869-30
9	43	60.6	15	9	US-09-798-869-7
10	42	59.2	15	9	US-09-798-869-4
11	42	59.2	25	9	US-09-798-869-22
12	39	54.9	15	9	US-09-798-869-8
13	39	54.9	338	9	US-09-798-295A-119
14	39	54.9	338	9	US-09-798-697-119
15	39	54.9	338	9	US-09-798-192A-119
16	39	54.9	338	9	US-09-999-832A-119
17	39	54.9	338	9	US-09-978-189-119
18	39	54.9	553	9	US-09-796-753-14
19	39	54.9	553	10	US-09-981-649A-6

20	39	54.9	553	10	US-09-981-649A-24	Sequence 24, Appl
21	39	54.9	554	10	US-09-981-649A-30	Sequence 30, Appl
22	39	54.9	554	10	US-09-981-649A-32	Sequence 32, Appl
23	39	54.9	559	10	US-09-981-649A-28	Sequence 28, Appl
24	38	53.5	333	9	US-09-796-753-26	Sequence 26, Appl
25	37	52.1	21	10	US-09-864-761-47985	Sequence 47985, A
26	36	50.7	13	9	US-09-798-869-26	Sequence 26, Appl
27	36	50.7	14	9	US-09-798-869-25	Sequence 25, Appl
28	36	50.7	15	9	US-09-798-869-1	Sequence 1, Appl
29	36	50.7	15	9	US-09-798-869-10	Sequence 10, Appl
30	36	50.7	15	9	US-09-798-869-28	Sequence 28, Appl
31	36	50.7	18	9	US-09-798-869-24	Sequence 24, Appl
32	36	50.7	25	9	US-09-798-869-21	Sequence 21, Appl
33	36	50.7	25	10	US-09-909-652-4	Sequence 4, Appl
34	36	50.7	25	10	US-09-030-619-205	Sequence 205, Appl
35	36	50.7	25	10	US-09-917-340-16	Sequence 16, Appl
36	35	50.7	846	9	US-10-051-409-4	Sequence 4, Appl
37	35	49.3	34	9	US-09-510-332-68	Sequence 68, Appl
38	35	49.3	61	10	US-09-864-761-40332	Sequence 40332, A
39	35	49.3	67	9	US-09-796-692-1487	Sequence 1487, Ap
40	35	49.3	489	9	US-09-888-320-2	Sequence 2, Appl
41	34	47.9	62	10	US-09-815-242-12129	Sequence 12129, A
42	34	47.9	62	10	US-09-815-242-13026	Sequence 13026, A
43	34	47.9	95	10	US-09-764-864-1031	Sequence 1031, Ap
44	34	47.9	351	10	US-09-853-625B-16	Sequence 16, Appl
45	34	47.9	747	9	US-10-066-500-58	Sequence 58, Appl

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 95.8%; Score 68; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQQRNMKKVR 12  
Db 3 CFQQRNMKKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON

```

; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-798-869-20

Query Match      95.8%; Score 68; DB 9; Length 25;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12
   |||||:||||
Db 3 CFQWQRMKKVR 14
   |||||:||||

RESULT 3
US-10-023-096-2
; Sequence 2, Application US/10023096
; Patent No. US20020160941A1
; GENERAL INFORMATION:
; APPLICANT: Kruzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Golinick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: 10505/P58185C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-10-023-096-2

```

```

Query Match      95.8%; Score 68; DB 9; Length 694;
Best Local Similarity 91.7%; Pred. No. 0.00088;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12
   |||||:||||
Db 22 CFQWQRMKKVR 33
   |||||:||||

RESULT 4
US-09-798-869-6
; Sequence 6, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens
; OTHER INFORMATION: sequence)
; US-09-798-869-6

Query Match      84.5%; Score 60; DB 9; Length 15;
Best Local Similarity 83.3%; Pred. No. 0.00051;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12
   |||||:||||
Db 3 CFQWQRMKKVR 14
   |||||:||||

RESULT 5
US-09-798-869-3
; Sequence 3, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: CAPRINE
; US-09-798-869-3

Query Match      71.8%; Score 51; DB 9; Length 15;
Best Local Similarity 63.6%; Pred. No. 0.015;

```

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
|:|||||:|:  
Db 3 CYQWQRMKKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 71.8%; Score 51; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.024;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
|:|||||:|:  
Db 3 CYQWQRMKKL 13

## RESULT 7

US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 63.4%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.15;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
|:|||||:|:  
Db 3 CFQWQRMKKL 13

## RESULT 8

US-09-798-869-30  
; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 63.4%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.15;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
|:|||||:|:  
Db 3 CFQWQRMKKL 13

## RESULT 9

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 60.6%; Score 43; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.31;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
|:|||||:|:  
Db 3 CFQWQRMKKL 13



RESULT 12  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:

```

RESULT 13
US - 09-978-295A-119
  / Sequence 119, Application US/09978295A
  / Patent No. US92020156006A1
  / GENERAL INFORMATION:
  / APPLICANT: Ashkenazi, Avi
  / APPLICANT: Baker Kevin P.
  / APPLICANT: Borstein, David
  / APPLICANT: Desnoyers, Luc
  / APPLICANT: Eaton, Dan
  / APPLICANT: Ferrara, Napoleon
  / APPLICANT: Filvaroff, Ellen
  / APPLICANT: Fong, Sherman
  / APPLICANT: Gao, Wei-Qiang
  / APPLICANT: Gerber, Hanspeter
  / APPLICANT: Gerritsen, Mary E.
  / APPLICANT: Goddard, Audrey
  / APPLICANT: Godowski, Paul J.
  / APPLICANT: Grimaldi, J. Christopher
  / APPLICANT: Gurney, Austin L.
  / APPLICANT: Hillan, Kenneth J
  / APPLICANT: Kljavin, Ivar J.
  / APPLICANT: Kuo, Sophia S.
  / APPLICANT: Napier, Mary A.
  / APPLICANT: Pan, James;
  / APPLICANT: Paoni, Nicholas F.
  / APPLICANT: Roy, Margaret Ann
  / APPLICANT: Shelton, David L.
  / APPLICANT: Stewart, Timothy A.
  / APPLICANT: Tunas, Daniel
  / APPLICANT: Williams, P. Mickey
  / APPLICANT: Wood, William I.
  / TITLE OF INVENTION: Acids and Trans
  / TITLE OF INVENTION: Acids Encoding th
  / FILE REFERENCE: P3630PlC11
  / CURRENT APPLICATION NUMBER: US/09/978
  / PRIOR FILING DATE: 2001-10-15
  / PRIOR APPLICATION NUMBER: 09/918585
  / PRIOR FILING DATE: 2001-07-30
  / PRIOR APPLICATION NUMBER: 60/062250
  / PRIOR FILING DATE: 1997-10-17

```

1 PRIOR APPLICATION NUMBER: 60/064249  
2 PRIOR FILING DATE: 1997-11-03  
3 PRIOR APPLICATION NUMBER: 60/065311  
4 PRIOR FILING DATE: 1997-11-13  
5 PRIOR APPLICATION NUMBER: 60/066364  
6 PRIOR FILING DATE: 1997-11-21  
7 PRIOR APPLICATION NUMBER: 60/077450  
8 PRIOR FILING DATE: 1998-03-10  
9 PRIOR APPLICATION NUMBER: 60/077632  
10 PRIOR FILING DATE: 1998-03-11  
11 PRIOR APPLICATION NUMBER: 60/077641  
12 PRIOR FILING DATE: 1998-03-11  
13 PRIOR APPLICATION NUMBER: 60/077649  
14 PRIOR FILING DATE: 1998-03-11  
15 PRIOR APPLICATION NUMBER: 60/077791  
16 PRIOR FILING DATE: 1998-03-12  
17 PRIOR APPLICATION NUMBER: 60/078004  
18 PRIOR FILING DATE: 1998-03-13  
19 PRIOR APPLICATION NUMBER: 60/078886  
20 PRIOR FILING DATE: 1998-03-20  
21 PRIOR APPLICATION NUMBER: 60/078936  
22 PRIOR FILING DATE: 1998-03-20  
23 PRIOR APPLICATION NUMBER: 60/078910  
24 PRIOR FILING DATE: 1998-03-20  
25 PRIOR APPLICATION NUMBER: 60/078939  
26 PRIOR FILING DATE: 1998-03-20  
27 PRIOR APPLICATION NUMBER: 60/079294  
28 PRIOR FILING DATE: 1998-03-25  
29 PRIOR APPLICATION NUMBER: 60/079656  
30 PRIOR FILING DATE: 1998-03-26  
31 PRIOR APPLICATION NUMBER: 60/079664  
32 PRIOR FILING DATE: 1998-03-27  
33 PRIOR APPLICATION NUMBER: 60/079689  
34 PRIOR FILING DATE: 1998-03-27  
35 PRIOR APPLICATION NUMBER: 60/079663  
36 PRIOR FILING DATE: 1998-03-27  
37 PRIOR APPLICATION NUMBER: 60/079728  
38 PRIOR FILING DATE: 1998-03-27  
39 PRIOR APPLICATION NUMBER: 60/079786  
40 PRIOR FILING DATE: 1998-03-27  
41 PRIOR APPLICATION NUMBER: 60/079920  
42 PRIOR FILING DATE: 1998-03-30  
43 PRIOR APPLICATION NUMBER: 60/079923  
44 PRIOR FILING DATE: 1998-03-30  
45 PRIOR APPLICATION NUMBER: 60/080105  
46 PRIOR FILING DATE: 1998-03-31  
47 PRIOR APPLICATION NUMBER: 60/080107  
48 PRIOR FILING DATE: 1998-03-31  
49 PRIOR APPLICATION NUMBER: 60/080165  
50 PRIOR FILING DATE: 1998-03-31  
51 PRIOR APPLICATION NUMBER: 60/080194  
52 PRIOR FILING DATE: 1998-03-31  
53 PRIOR APPLICATION NUMBER: 60/080327  
54 PRIOR FILING DATE: 1998-04-01  
55 PRIOR APPLICATION NUMBER: 60/080328  
56 PRIOR FILING DATE: 1998-04-01  
57 PRIOR APPLICATION NUMBER: 60/080333  
58 PRIOR FILING DATE: 1998-04-01  
59 PRIOR APPLICATION NUMBER: 60/080334  
60 PRIOR FILING DATE: 1998-04-01  
61 PRIOR APPLICATION NUMBER: 60/080328  
62 PRIOR FILING DATE: 1998-04-01  
63 PRIOR APPLICATION NUMBER: 60/081049  
64 PRIOR FILING DATE: 1998-04-08  
65 PRIOR APPLICATION NUMBER: 60/081071  
66 PRIOR FILING DATE: 1998-04-08  
67 PRIOR APPLICATION NUMBER: 60/081195  
68 PRIOR FILING DATE: 1998-04-08  
69 PRIOR APPLICATION NUMBER: 60/081203  
70 PRIOR FILING DATE: 1998-04-09  
71 PRIOR APPLICATION NUMBER: 60/081229  
72 PRIOR FILING DATE: 1998-04-09  
73 PRIOR APPLICATION NUMBER: 60/081955

74 PRIOR FILING DATE: 1998-04-15  
75 PRIOR APPLICATION NUMBER: 60/081817  
76 PRIOR FILING DATE: 1998-04-15  
77 PRIOR APPLICATION NUMBER: 60/081819  
78 PRIOR FILING DATE: 1998-04-15  
79 PRIOR APPLICATION NUMBER: 60/081952  
80 PRIOR FILING DATE: 1998-04-15  
81 PRIOR APPLICATION NUMBER: 60/081838  
82 PRIOR FILING DATE: 1998-04-15  
83 PRIOR APPLICATION NUMBER: 60/082568  
84 PRIOR FILING DATE: 1998-04-21  
85 PRIOR APPLICATION NUMBER: 60/082569  
86 PRIOR FILING DATE: 1998-04-21  
87 PRIOR APPLICATION NUMBER: 60/082704  
88 PRIOR FILING DATE: 1998-04-22  
89 PRIOR APPLICATION NUMBER: 60/082804  
90 PRIOR FILING DATE: 1998-04-22  
91 PRIOR APPLICATION NUMBER: 60/082700  
92 PRIOR FILING DATE: 1998-04-22  
93 PRIOR APPLICATION NUMBER: 60/082797  
94 PRIOR FILING DATE: 1998-04-22  
95 PRIOR APPLICATION NUMBER: 60/082796  
96 PRIOR FILING DATE: 1998-04-23  
97 PRIOR APPLICATION NUMBER: 60/083336  
98 PRIOR FILING DATE: 1998-04-27  
99 PRIOR APPLICATION NUMBER: 60/083322  
100 PRIOR FILING DATE: 1998-04-28  
101 PRIOR APPLICATION NUMBER: 60/083392  
102 PRIOR FILING DATE: 1998-04-29  
103 PRIOR APPLICATION NUMBER: 60/083495  
104 PRIOR FILING DATE: 1998-04-29  
105 PRIOR APPLICATION NUMBER: 60/083496  
106 PRIOR FILING DATE: 1998-04-29  
107 PRIOR APPLICATION NUMBER: 60/083499  
108 PRIOR FILING DATE: 1998-04-29  
109 PRIOR APPLICATION NUMBER: 60/083545  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083554  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083558  
114 PRIOR FILING DATE: 1998-04-29  
115 PRIOR APPLICATION NUMBER: 60/083559  
116 PRIOR FILING DATE: 1998-04-29  
117 PRIOR APPLICATION NUMBER: 60/083500  
118 PRIOR FILING DATE: 1998-04-29  
119 PRIOR APPLICATION NUMBER: 60/083742  
120 PRIOR FILING DATE: 1998-04-30  
121 PRIOR APPLICATION NUMBER: 60/084366  
122 PRIOR FILING DATE: 1998-05-05  
123 PRIOR APPLICATION NUMBER: 60/084414  
124 PRIOR FILING DATE: 1998-05-06  
125 PRIOR APPLICATION NUMBER: 60/084441  
126 PRIOR FILING DATE: 1998-05-06  
127 PRIOR APPLICATION NUMBER: 60/084637  
128 PRIOR FILING DATE: 1998-05-07  
129 PRIOR APPLICATION NUMBER: 60/084639  
130 PRIOR FILING DATE: 1998-05-07  
131 PRIOR APPLICATION NUMBER: 60/084640  
132 PRIOR FILING DATE: 1998-05-07  
133 PRIOR APPLICATION NUMBER: 60/084598  
134 PRIOR FILING DATE: 1998-05-07  
135 PRIOR APPLICATION NUMBER: 60/084600  
136 PRIOR FILING DATE: 1998-05-07  
137 PRIOR APPLICATION NUMBER: 60/084627  
138 PRIOR FILING DATE: 1998-05-07  
139 PRIOR APPLICATION NUMBER: 60/084643  
140 PRIOR FILING DATE: 1998-05-07  
141 PRIOR APPLICATION NUMBER: 60/085339  
142 PRIOR FILING DATE: 1998-05-13  
143 PRIOR APPLICATION NUMBER: 60/085338  
144 PRIOR FILING DATE: 1998-05-13  
145 PRIOR APPLICATION NUMBER: 60/085323  
146 PRIOR FILING DATE: 1998-05-13

;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 54.9%; Score 39; DB 9; Length 338;

Best Local Similarity 54.5%; Pred. No. 26;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQORNMKV 11

Db 50 CYGWRNSKGV 60

#### RESULT 14

US-09-978-697-119

Sequence 119, Application US/09978697

Patent No. US20020169284A1

GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kijavrin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; TITLE OF INVENTION: Acids Encoding the Same  
;; FILE REFERENCE: P2630PIC27  
;; CURRENT APPLICATION NUMBER: US/09/978,697  
;; CURRENT FILING DATE: 2001-10-16  
;; PRIOR APPLICATION NUMBER: 09/918585  
;; PRIOR FILING DATE: 2001-07-30  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064249  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/077450

;; PRIOR FILING DATE: 1998-03-10  
;; PRIOR APPLICATION NUMBER: 60/077632  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077641  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077649  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078004  
;; PRIOR FILING DATE: 1998-03-13  
;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15

;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579

;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 54.9%; Score 39; DB 9; Length 338;  
Best Local Similarity 54.5%; Pred. No. 26;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNMKKV 11  
|:|:|:|:|  
Db 50 CYGWRNRSGV 60

## RESULT 15

US-09-978-192A-119  
; Sequence 119, Application US/09978192A  
; Patent No. US2002017753A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kijavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James,  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC9  
; CURRENT APPLICATION NUMBER: US/09/978,192A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11

1	PRIOR APPLICATION NUMBER: 60/077919	
2	PRIOR FILING DATE: 1998-03-12	
3	PRIOR APPLICATION NUMBER: 60/078004	
4	PRIOR FILING DATE: 1998-03-13	
5	PRIOR APPLICATION NUMBER: 60/078886	
6	PRIOR FILING DATE: 1998-03-20	
7	PRIOR APPLICATION NUMBER: 60/078936	
8	PRIOR FILING DATE: 1998-03-20	
9	PRIOR APPLICATION NUMBER: 60/078910	
10	PRIOR FILING DATE: 1998-03-20	
11	PRIOR APPLICATION NUMBER: 60/078939	
12	PRIOR FILING DATE: 1998-03-20	
13	PRIOR APPLICATION NUMBER: 60/079294	
14	PRIOR FILING DATE: 1998-03-25	
15	PRIOR APPLICATION NUMBER: 60/079656	
16	PRIOR FILING DATE: 1998-03-26	
17	PRIOR APPLICATION NUMBER: 60/079664	
18	PRIOR FILING DATE: 1998-03-27	
19	PRIOR APPLICATION NUMBER: 60/079668	
20	PRIOR FILING DATE: 1998-03-27	
21	PRIOR APPLICATION NUMBER: 60/079663	
22	PRIOR FILING DATE: 1998-03-27	
23	PRIOR APPLICATION NUMBER: 60/079728	
24	PRIOR FILING DATE: 1998-03-27	
25	PRIOR APPLICATION NUMBER: 60/079786	
26	PRIOR FILING DATE: 1998-03-27	
27	PRIOR APPLICATION NUMBER: 60/079920	
28	PRIOR FILING DATE: 1998-03-30	
29	PRIOR APPLICATION NUMBER: 60/079923	
30	PRIOR FILING DATE: 1998-03-30	
31	PRIOR APPLICATION NUMBER: 60/080105	
32	PRIOR FILING DATE: 1998-03-31	
33	PRIOR APPLICATION NUMBER: 60/080107	
34	PRIOR FILING DATE: 1998-03-31	
35	PRIOR APPLICATION NUMBER: 60/080165	
36	PRIOR FILING DATE: 1998-03-31	
37	PRIOR APPLICATION NUMBER: 60/080194	
38	PRIOR FILING DATE: 1998-03-31	
39	PRIOR APPLICATION NUMBER: 60/080327	
40	PRIOR FILING DATE: 1998-04-01	
41	PRIOR APPLICATION NUMBER: 60/080328	
42	PRIOR FILING DATE: 1998-04-01	
43	PRIOR APPLICATION NUMBER: 60/080333	
44	PRIOR FILING DATE: 1998-04-01	
45	PRIOR APPLICATION NUMBER: 60/080334	
46	PRIOR FILING DATE: 1998-04-01	
47	PRIOR APPLICATION NUMBER: 60/081070	
48	PRIOR FILING DATE: 1998-04-08	
49	PRIOR APPLICATION NUMBER: 60/081049	
50	PRIOR FILING DATE: 1998-04-08	
51	PRIOR APPLICATION NUMBER: 60/081071	
52	PRIOR FILING DATE: 1998-04-08	
53	PRIOR APPLICATION NUMBER: 60/081195	
54	PRIOR FILING DATE: 1998-04-08	
55	PRIOR APPLICATION NUMBER: 60/081203	
56	PRIOR FILING DATE: 1998-04-09	
57	PRIOR APPLICATION NUMBER: 60/081229	
58	PRIOR FILING DATE: 1998-04-09	
59	PRIOR APPLICATION NUMBER: 60/081955	
60	PRIOR FILING DATE: 1998-04-15	
61	PRIOR APPLICATION NUMBER: 60/081817	
62	PRIOR FILING DATE: 1998-04-15	
63	PRIOR APPLICATION NUMBER: 60/081819	
64	PRIOR FILING DATE: 1998-04-15	
65	PRIOR APPLICATION NUMBER: 60/081952	
66	PRIOR FILING DATE: 1998-04-15	
67	PRIOR APPLICATION NUMBER: 60/081838	
68	PRIOR FILING DATE: 1998-04-15	
69	PRIOR APPLICATION NUMBER: 60/082568	
70	PRIOR FILING DATE: 1998-04-21	
71	PRIOR APPLICATION NUMBER: 60/082569	
72	PRIOR FILING DATE: 1998-04-21	
73	PRIOR APPLICATION NUMBER: 60/082704	

; PRIOR APPLICATION NUMBER: 60/085697

Query Match 54.9%; Score 39; DB 9; Length 338;  
Best Local Similarity 54.5%; Pred. No. 26;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWORNKKV 11  
|:|:|:|  
Db 50 CYGWRNRSKV 60

Search completed: February 21, 2003, 08:08:08  
Job time : 10.55 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-83

Perfect score: 71

Sequence: 1 CFQWQNNKKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	95.8	711	1 TFHUL	lactotransferrin p
2	51	71.8	708	2 JC2323	lactoferrin - goat
3	48	67.6	33	2 S52107	lactoferrin - sheep
4	42	59.2	511	2 A80858	hypothetical prote
5	42	59.2	707	1 A28438	lactoferrin precu
6	41	57.7	282	2 F90580	hypothetical prote
7	41	57.7	4568	2 T08030	dynein beta heavy
8	40	56.3	275	1 JC1113	interleukin-2 rece
9	40	56.3	275	1 S07442	interleukin-2 rece
10	40	56.3	584	2 C84325	hypothetical prote
11	39	54.9	205	2 E90094	26S proteasome SU
12	39	54.9	298	2 A02346	hypothetical prote
13	39	54.9	531	2 A84471	En/spn-like transp
14	39	54.9	558	2 T17324	hypothetical prote
15	38	53.5	223	2 T37974	probable peroxisom
16	38	53.5	238	2 T40568	hypothetical prote
17	38	53.5	275	2 T22597	hypothetical prote
18	38	53.5	323	2 C92234	conserved hypotet
19	38	53.5	335	2 T33211	hypothetical prote
20	38	53.5	393	2 D71876	3-deoxy-manno-octu
21	38	53.5	515	2 T00510	probable cytochrom
22	38	53.5	543	2 T00513	cycochrome P450 ho
23	38	53.5	932	2 T28820	hypothetical prote
24	38	53.5	1436	2 F86904	conserved hypotet
25	37	52.1	99	2 C97838	hypothetical prote
26	37	52.1	283	2 D72378	sugar ABC transpor
27	37	52.1	376	2 S67085	hypothetical prote
28	37	52.1	393	2 E64639	3-deoxy-manno-octu
29	37	52.1	464	2 A12343	hypothetical prote

## ALIGNMENTS

### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74:

R/Cho, Y.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>

A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R/Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>

A/Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactoferi

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <TEN>

A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A/Experimental source: placenta

A/Note: sequence extracted from NCBI backbone (NCBIP:122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POW>

A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 20-31 <ST1>

A/Accession: S20841

A/Molecule type: protein

A/Residues: 20-28, 'X', 30-31 <ST2>



```

Query Match          71.8%; Score 51; DB 2; Length 700;
Best Local Similarity 63.6%; Pred. No. 0.44;
Matches 7; Conservative 3; Mismatches 1; Indels

QY 1 CFQWQRNMKKV 11
   ||||| ||:
DB 38 CYQWRMRKL 48

RESULT 3
S52107
lactoferrin - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sh
C:Date: 14-Jul-1995 #sequence_revision 05-Dec-1997 #text_change
C:Accession: S52107
Biochim. Biophys. Acta 1243, 25-32, 1995
R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.
A:Title: Isolation and characterization of sheep lactoferrin, a
A:Reference number: S52107; NCID:95127729; PMID:7827104
A:Accession: S52107
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-33 <QIA>
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication

Query Match          67.6%; Score 48; DB 2; Length 33;
Best Local Similarity 54.5%; Pred. No. 0.07;
Matches 6; Conservative 4; Mismatches 1; Indels

QY 1 CFQWQRNMKKV 11
   ||||| ||:
DB 19 CYQWRMRKL 29

RESULT 4
AB0858
hypothetical protein STY3070 [imported] - Salmonella enterica s
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change
C:Accession: AB0858
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard
th, T.; Comerton, F.; Cronin, A.; Davis, P.; Davies, R.M.; Dow
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; S
A:Title: Complete genome sequence of a multiple drug resistant
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0858
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:g16504016
C:Genetics:
A:Gene: STY3070

Query Match          59.2%; Score 42; DB 2; Length 511;
Best Local Similarity 58.3%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 5; Indels

QY 1 CFQWQRNMKKVR 12
   ||||| ||||
DB 350 CFQWDMNKKVR 361

RESULT 5
A28438
lactoferrin precursor - mouse
N:Alternate names: lactotransferrin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change

```

C;Accession: A28438; A41205  
 R;Pentecost, B. T.; Teng, C. T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory epithelium  
 A;Reference number: A92596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Lin, Y.; Teng, C. T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) #status predicted

Query Match 59.2%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 17;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWRNKKV 11  
 | : : : : :  
 Db 37 CLRQWQNRKV 47

RESULT 6  
 F90580  
 Hypothetical protein MYPV 5500 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C;Species: Mycoplasma pulmonis  
 C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C;Accession: F90580  
 R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001  
 A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis  
 A;Reference number: A99512; MUID:21267165; PMID:11353084  
 A;Accession: F90580  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-282 <KUR>  
 A;Cross-references: GB:AL445566; PID:gl4089965; PIDN:CAC13723.1; GSPDB:GN00153  
 A;Experimental source: strain UAB CTIP  
 C;Genetics:  
 A;Gene: MYPV 5500  
 A;Genetic code: SGC3

Query Match 57.7%; Score 41; DB 2; Length 282;  
 Best Local Similarity 60.0%; Pred. No. 10;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWRNKKV 11  
 | : : : : :  
 Db 20 FAWQNNIKKI 29

RESULT 7  
 T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C;Species: Chlamydomonas reinhardtii  
 C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 C;Accession: T08030  
 R;Mitchell, D. R.; Brown, K. S.  
 J. Cell Sci. 107, 635-644, 1994  
 A;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
 A;Reference number: T16302; MUID:94274778; PMID:8006077  
 A;Accession: T08030

A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-4568 <MT>  
 A;Cross-references: EMBL:U02963; NID:9409965; PIDN:AAAI9956.1; PID:9514215  
 A;Experimental source: strain 21gr  
 C;Genetics:  
 A;Gene: ODA4  
 A;Map position: IX  
 A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3; 1334/3; 3686/3; 3882/3; 4240/3  
 C;Superfamily: dynein heavy chain, ciliary  
 C;Keywords: nucleotide binding; P-loop  
 F;1919-1926/Region: nucleotide-binding motif A (P-loop)  
 F;2202-2209/Region: nucleotide-binding motif A (P-loop)  
 F;2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 57.7%; Score 41; DB 2; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWRNKKV 12  
 | : : : : :  
 Db 1852 CFQWQSLRYIQ 1863

RESULT 8  
 JC1113  
 Interleukin-2 receptor alpha chain precursor - sheep  
 N;Alternate names: CD25  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 11-Sep-1998 #sequence\_revision 11-Sep-1998 #text\_change 22-Jun-1999  
 C;Accession: JC1113; S18899; S18910  
 R;Bujdosó, R.; Sargan, D.; Williamson, M.; McConnell, I.  
 Gene 113, 283-284, 1992  
 A;Title: Cloning of a cDNA encoding the ovine interleukin-2 receptor 55-kDa protein, CD25  
 A;Reference number: JC1113; MUID:92241682; PMID:1572550  
 A;Accession: JC1113  
 A;Molecule type: mRNA  
 A;Residues: 1-275 <BUJ>  
 A;Cross-references: EMBL:X60149; NID:gl1287; PIDN:CAA42723.1; PID:gl1288  
 R;Verhagen, A. A.  
 submitted to the EMBL Data Library, December 1991  
 A;Description: Molecular cloning, expression and characterisation of the ovine IL-2R alpha  
 A;Reference number: S18899  
 A;Accession: S18899  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-165, 'S', 167-275 <VER>  
 A;Cross-references: EMBL:Z11560; NID:gl1275; PIDN:CAA77652.1; PID:gl1276  
 C;Complex: The high affinity receptor is a heterotrimer of alpha, beta, and gamma chains;  
 C;Function:  
 A;Description: receptor for interleukin-2  
 A;Pathway: interleukin-2 stimulated growth and differentiation of T cells, B cells, NK cells  
 C;Superfamily: interleukin-2 receptor alpha chain; complement factor H repeat homology  
 C;Keywords: cytokine receptor; duplication; glycoprotein; T-cell proliferation; transmembrane  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-275/Product: interleukin-2 receptor alpha chain #status predicted <MAT>  
 F;22-243/Domain: extracellular #status predicted <EXT>  
 F;24-77/Domain: complement factor H repeat homology <FH1>  
 F;123-184/Domain: complement factor H repeat homology <FH2>  
 F;244-264/Domain: transmembrane #status predicted <TM>  
 F;265-275/Domain: intracellular #status predicted <INT>  
 F;24-64, 51-77, 123-168, 152-184/Disulfide bonds: #status predicted  
 F;80/Binding site: carbohydrate (Asn) #status predicted

Query Match 56.3%; Score 40; DB 1; Length 275;  
 Best Local Similarity 58.3%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWRNKKV 12  
 | : : : : :  
 Db 261 CLATWRNKKNR 272

RESULT 9  
S07442  
interleukin-2 receptor alpha chain precursor - bovine  
N;Alternate names: CD25  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 11-Sep-1998 #sequence\_revision 11-Sep-1998 #text\_change 22-Jun-1999  
C;Accession: S07442  
R;Weinberg, A.D.; Shaw, J.; Paetkau, V.; Bleackley, R.C.; Magnuson, N.S.; Reeves, R.; Ma  
Immunology 63, 603-610, 1988  
A;Title: Cloning of cDNA for the bovine IL-2 receptor (bovine Tac antigen).  
A;Reference number: S07442; MUID:88212503; PMID:2835311  
A;Accession: S07442  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-275 <WEI>  
A;Cross-references: EMBL:M20818; NID:gl63208; PIDN:AA51414.1; PID:gl63209  
C;Complex: The high affinity receptor is a heterotrimer of alpha, beta, and gamma chains  
C;Function:  
A;Description: receptor for interleukin-2  
A;Pathway: interleukin-2 stimulated growth and differentiation of T cells, B cells, NK c  
C;Superfamily: interleukin-2 receptor alpha chain; complement factor H repeat homology  
C;Keywords: cytokine receptor; duplication; glycoprotein; T-cell proliferation; transmem  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-275/Product: interleukin-2 receptor alpha chain #status predicted <MAT>  
F;22-243/Domain: extracellular #status predicted <EXT>  
F;24-77/Domain: complement factor H repeat homology <FHL>  
F;123-184/Domain: complement factor H repeat homology <FH2>  
F;244-264/Domain: transmembrane #status predicted <TMM>  
F;265-275/Domain: intracellular #status predicted <INT>  
F;24-64,51-77,123-168,152-184/Disulfide bonds: #status predicted  
F;80,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 56.3%; Score 40; DB 1; Length 275;  
Best Local Similarity 58.3%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
| | | | |  
Db 261 CLTWQKWKNNR 272

RESULT 10  
C84325  
hypothetical protein Vngl732c [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: C84325  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabc  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: C84325  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-584 <STO>  
A;Cross-references: GB:AE004437; NID:gl0581192; PIDN:AAG19967.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: VNG1732C

Query Match 56.3%; Score 40; DB 2; Length 584;  
Best Local Similarity 41.7%; Pred. No. 32;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
| | | | |  
Db 445 CFTWRKDKMERK 456

RESULT 11

E90094  
26S proteasome SU B5 [imported] - Guillardia theta nucleomorph  
C;Species: nucleomorph Guillardia theta  
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C;Accession: E90094  
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei  
Nature 410, 1091-1096, 2001  
A;Title: The highly reduced genome of an enslaved algal nucleus.  
A;Reference number: A99082; MUID:11323671; PMID:11323671  
A;Accession: E90094  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-205 <DOU>  
A;Cross-references: GB:AF165818; NID:gl3794510; PIDN:AAK39885.1; GSPDB:GN00150  
C;Genetics:  
A;Gene: prsB5  
A;Map position: 1  
A;Genome: nucleomorph  
C;Keywords: nucleomorph

Query Match 54.9%; Score 39; DB 2; Length 205;  
Best Local Similarity 45.5%; Pred. No. 17;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQNNMKV 11  
| | | | |  
Db 63 CFFWERNLSSL 73

RESULT 12  
AD2346  
hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp.  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C;Accession: AD2346  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AD2346  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-298 <KOR>  
A;Cross-references: GB:BA000019; PIDN:BA876022.1; PID:gl7133459; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr4323

Query Match 54.9%; Score 39; DB 2; Length 298;  
Best Local Similarity 66.7%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQNNMKK 10  
| | | | |  
Db 163 FHWQNNYRK 171

RESULT 13  
A84471  
En/Spm-like transposon protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: A84471  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: A84471

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-531 <STO>  
A;Cross-references: GB:AE002093; NID:G4586022; PIDN:AAD25641.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g05650  
A;Map position: 2

Search completed: February 21, 2003, 07:47:55  
Job time : 10.65 secs

Query Match 54.9%; Score 39; DB 2; Length 531;  
Best Local Similarity 87.5%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QWQRNMKK 10  
|||  
Db 501 QWFRNMKK 508

## RESULT 14

Tl7324  
hypothetical protein DKFZp564P2063.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: Tl7324  
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A;Reference number: Z18727  
A;Accession: Tl7324  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-558 <DUE>  
A;Cross-references: EMBL:AL117610  
A;Experimental source: fetal brain; clone DKFZp564P2063  
C;Genetics:  
A;Note: DKFZp564P2063.1

Query Match 54.9%; Score 39; DB 2; Length 558;  
Best Local Similarity 54.5%; Pred. No. 46;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWRNMKKV 11  
|:|:|  
Db 55 CYGWRNRSKGV 65

## RESULT 15

T37974  
probable peroxisomal membrane protein - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
C;Accession: T37974  
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z21759  
A;Accession: T37974  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-223 <MUR>  
A;Cross-references: EMBL:Z69909; PIDN:CAA93785.1; GSPDB:GN00066; SPDB:SPAC19G10.03c  
A;Experimental source: strain 972h-; cosmid cl9G10  
C;Genetics:  
A;Gene: SPDB:SPAC19G10.03c  
A;Map position: 1  
A;Introns: 10/3; 170/2  
C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC582.09  
C;Keywords: peroxisome

Query Match 53.5%; Score 38; DB 2; Length 223;  
Best Local Similarity 41.7%; Pred. No. 28;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWRNMKKVR 12  
|:|:|  
Db 140 CYELQONSKKIK 151

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 Seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-83  
Perfect score: 71  
Sequence: 1 CFQWQNMKKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	711	1 TRFL_HUMAN	P02788 homo sapien
2	57	80.3	708	1 TRFL_CAMDR	Q9tum0 camelus dro
3	51	71.8	708	1 TRFL_CAPHI	Q29477 capra hircu
4	49	69.0	695	1 TRFL_HORSE	O77811 equus cabal
5	42	59.2	707	1 TRFL_MOUSE	P08071 mus musculu
6	41	57.7	4568	1 DYHB_CHLRE	Q39565 chlamydomon
7	40	56.3	275	1 IL2A_BOVIN	P12342 bos taurus
8	40	56.3	275	1 IL2A_SHEEP	P26898 ovis aries
9	38	53.5	238	1 YBM9_SCHPO	Q10333 schizosacch
10	38	53.5	708	1 TRFL_BUBBU	O77698 bubalus bub
11	37	52.1	453	1 O83A_DROME	Q9vnb3 drosophila
12	37	52.1	749	1 VP4_ROTGA	Q04916 rotavirus (
13	36.5	51.4	425	1 PTA_HAEIN	P45068 haemophilus
14	36	50.7	275	1 VNS2_DSDNV	O71154 diatraea sa
15	36	50.7	292	1 NLA_DROME	Q9x218 drosophila
16	36	50.7	329	1 CATK_RAT	Q35186 rattus norv
17	36	50.7	435	1 DCOR_PANRE	P49725 panagrellus
18	36	50.7	665	1 YL14_CABEL	Q11100 canorhabdi
19	36	50.7	708	1 TRFL_BOVIN	P24527 bos taurus
20	36	50.7	765	1 Y008_HUMAN	Q15398 homo sapien
21	36	50.7	989	1 T100_HUMAN	O75448 homo sapien
22	36	50.7	1135	1 PHYC_SORBI	P93528 sorghum bic
23	36	50.7	2594	1 7LES_DROVI	P20806 drosophila
24	35	49.3	146	1 RPOB_LIBAF	P41187 liberibacte
25	35	49.3	160	1 Y4LN_RHISN	P55554 rhizobium s
26	35	49.3	211	1 LOLE_VIBCH	P57070 vibrio chol
27	35	49.3	485	1 GLGA_BACST	O08328 bacillus st
28	35	49.3	502	1 C911_ARATH	Q9f965 arabidopsis
29	35	49.3	528	1 CAX2_ARATH	Q38798 arabidopsis
30	35	49.3	530	1 CAX1_ARATH	P29402 arabidopsis
31	35	49.3	569	1 Y397_MYCPN	Q50333 mycoplasma
32	35	49.3	573	1 PREL_LACFE	P26929 lactobacill
33	35	49.3	612	1 RPSD_BUCAI	P57163 buchnera ap

## RESULT 1

ID	TRFL_HUMAN	STANDARD;	PRT;	711 AA.
AC	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9HLZ3; Q96KZ4;			
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor [Lactoferrin] [Contains: Lactoferrin A;			
DE	Lactoferrin B; Lactoferrin C].			
GN	LTF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cho Y.Y.;			
RL	Thesis (1994), Genetic Engineering Research Institute / Taejeon, Korea.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Connely O.M.;			
RT	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RT	"Molecular cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RL	sequences.";			
RN	[6]			
RP	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RT	"cDNA cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[8]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Straussberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			

## ALIGNMENTS

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=8507667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RT Legrand D., Spik G., Montreuil J., Jolles P.;  
RL "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gao J.L., la Baidie M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Segripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253--methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmanik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
RN [21]  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -!- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53961; CAA37914.1; -;  
DR EMBL; U07643; AAB60324.1; -;  
DR EMBL; M93150; AAA36159.1; -;  
DR EMBL; M83202; AAA59511.1; -;  
DR EMBL; M83205; AAA58656.1; -;  
DR EMBL; M18642; AAA86665.1; -;  
DR EMBL; AF332168; AAG48753.1; -;  
DR EMBL; BC015822; AAH15822.1; -;  
DR EMBL; BC015823; AAH15823.1; -;  
DR EMBL; M73700; AAA59479.1; -;  
DR EMBL; X52941; CAA37116.1; -;  
DR EMBL; U95626; AAB57795.1; -;  
DR PIR; S11228; TFHUL.  
DR PDB; 1LCF; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LPH; 31-OCT-93.  
DR PDB; 1LFI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1BKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

```

Query Match          95.8%; Score 68; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.00022;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWQRNMKVR 12
   |||||:|:|
Db 39 CFQWQRNMKVR 50

RESULT 2
ID TRFL CAMDR STANDARD; PRT; 708 AA.
AC Q9TUMG; Q9MZS5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactoferrin precursor (Lactoferrin).
GN LTF.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Somali; TISSUE=Lactating mammary gland;
RA Kappeler S.R., Ackermann M., Parah Z., Puhon Z.;
RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";
RL Int. Dairy J. 9:481-486(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AJ131674; CAB53387.1; -
DR EMBL: AF165879; AAF82241.1; -
DR HSPG; O77811; 1b1x.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SMC00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN 1; 2.
DR PROSITE; PS00206; TRANSFERRIN 2; 2.
DR PROSITE; PS00207; TRANSFERRIN 3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
Signal.
FT CHAIN 1 19 BY SIMILARITY.
FT REPEAT 20 708 LACTOTRANSFERRIN.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.

```

```

FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 261 261 F -> S (IN REF. 2).
FT CONFLICT 304 304 G -> A (IN REF. 2).
FT CONFLICT 330 330 S -> P (IN REF. 2).
FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).
FT CONFLICT 506 506 L -> F (IN REF. 2).
FT CONFLICT 609 609 A -> P (IN REF. 2).
FT CONFLICT 642 642 R -> Q (IN REF. 2).
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match      80.3%; Score 57; DB 1; Length 708;
Best Local Similarity 83.3%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQRNMKVR 12
   |||||:|:|
Db 38 CAQWQRNMKVR 49

RESULT 3
ID TRFL CAPHI STANDARD; PRT; 708 AA.
AC Q29477; Q29479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Lee T., Yu S., Kim S., Lee K., Yu D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=94380047; PubMed=8093048;
RA le Provost F., Nocard M., Guerin G., Martin P.;
RT "Characterization of the goat lactoferrin cDNA. Assignment of the
RT relevant locus to bovine U12 syntenic group.";
RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U53857; AAA97958.1; -.
CC EMBL; X78902; CAA55517.1; -.
CC HSP; 077698; 1CE2.
CC InterPro; IPR001156; Transferrin.
CC Pfam; PF00405; transferrin; 2.
CC PRINTS; PR00422; TRANSFERRIN.
CC SMART; SM00094; TR_FER; 2.
CC PROSITE; PS00205; TRANSFERRIN_1; 2.
CC PROSITE; PS00206; TRANSFERRIN_2; 2.
CC PROSITE; PS00207; TRANSFERRIN_3; 2.
CC Transprot; Iron transport; Glycoprotein; Metal-binding; Repeat;
CC Signal.
CC FT SIGNAL 1 19 BY SIMILARITY.
CC FT CHAIN 20 708 LACTOTRANSFERRIN.
CC FT REPEAT 20 363 1.
CC FT REPEAT 364 708 2.
CC FT DISULFID 28 64 BY SIMILARITY.
CC FT DISULFID 38 55 BY SIMILARITY.
CC FT DISULFID 134 217 BY SIMILARITY.
CC FT DISULFID 176 192 BY SIMILARITY.
CC FT DISULFID 189 200 BY SIMILARITY.
CC FT DISULFID 250 264 BY SIMILARITY.
CC FT DISULFID 367 399 BY SIMILARITY.
CC FT DISULFID 377 390 BY SIMILARITY.
CC FT DISULFID 424 703 BY SIMILARITY.
CC FT DISULFID 444 666 BY SIMILARITY.
CC FT DISULFID 476 551 BY SIMILARITY.
CC FT DISULFID 500 694 BY SIMILARITY.
CC FT DISULFID 510 524 BY SIMILARITY.
CC FT DISULFID 521 534 BY SIMILARITY.
CC FT DISULFID 592 606 BY SIMILARITY.
CC FT DISULFID 644 649 BY SIMILARITY.
CC FT METAL 79 79 IRON 1 (BY SIMILARITY).
CC FT METAL 111 111 IRON 1 (BY SIMILARITY).
CC FT METAL 211 211 IRON 1 (BY SIMILARITY).
CC FT METAL 272 272 IRON 1 (BY SIMILARITY).
CC FT METAL 414 414 IRON 2 (BY SIMILARITY).
CC FT METAL 452 452 IRON 2 (BY SIMILARITY).
CC FT METAL 545 545 IRON 2 (BY SIMILARITY).
CC FT METAL 614 614 IRON 2 (BY SIMILARITY).
CC FT BINDING 140 140 ANION (BY SIMILARITY).
CC FT BINDING 482 482 ANION (BY SIMILARITY).
CC FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 56 56 I -> V (IN REF. 2).
CC FT CONFLICT 88 88 L -> R (IN REF. 2).
CC FT CONFLICT 124 124 Q -> K (IN REF. 2).
CC FT CONFLICT 154 154 P -> P (IN REF. 2).
CC FT CONFLICT 304 304 S -> R (IN REF. 2).
CC FT CONFLICT 414 414 D -> G (IN REF. 2).
CC SEQUENCE 708 AA; 77358 MW; F2ED3C83539960D CRC64;
Query Match 71.8%; Score 51; DB 1; Length 708;
Best Local Similarity 63.6%; Pred. No. 0.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CFQWQNNMKV 11
Db 38 CYQWQNNMKV 48

```

## RESULT 4

```

TRFL HORSE
ID TRFL_HORSE STANDARD; PRT; 695 AA.
AC 07811;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (lactoferrin) (fragment).
GN LTF.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "cDNA sequence of mare lactoferrin.";
RN [2] Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RC TISSUE=Milk;
RX MEDLINE=99236631; PubMed=10366507;
RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A
resolution.";
RL J. Mol. Biol. 289:303-317(1999).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ010930; CAA09407.1; -.
CC PDB; 1B1X; 02-DEC-98.
CC PDB; 1B7U; 02-FEB-99.
CC PDB; 1B7Z; 02-FEB-99.
CC InterPro; IPR001156; Transferrin.
CC Pfam; PF00405; transferrin; 2.
CC PRINTS; PR00422; TRANSFERRIN.
CC SMART; SM00094; TR_FER; 2.
CC PROSITE; PS00205; TRANSFERRIN_1; 2.
CC PROSITE; PS00206; TRANSFERRIN_2; 2.
CC PROSITE; PS00207; TRANSFERRIN_3; 1.
CC Transprot; Iron transport; Glycoprotein; Metal-binding; Repeat;
CC Signal; 3D-structure.
CC NON TER 1 1
CC SIGNAL <1 6 LACTOTRANSFERRIN.
CC FT CHAIN 7 695 1.
CC FT REPEAT 7 350 1.
CC FT REPEAT 351 695 2.
CC FT DISULFID 15 51
CC FT DISULFID 25 42
CC FT DISULFID 121 204
CC FT DISULFID 163 179
CC FT DISULFID 166 189
CC FT DISULFID 176 187
CC FT DISULFID 237 251
CC FT DISULFID 354 385
CC FT DISULFID 364 377
CC FT DISULFID 411 690
CC FT DISULFID 431 653

```



FT DISULFID 463 538  
 FT DISULFID 487 691  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66 66  
 FT METAL 98 98  
 FT METAL 198 198  
 FT METAL 259 259  
 FT METAL 401 401  
 FT METAL 439 439  
 FT METAL 532 532  
 FT METAL 601 601  
 FT BINDING 127 127  
 FT BINDING 469 469  
 FT CARBOHYD 143 143  
 FT CARBOHYD 287 287  
 FT CARBOHYD 482 482  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 69.0%; Score 49; DB 1; Length 695;  
 Best Local Similarity 75.0%; Pred. No. 0.44;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKKVR 12  
 DB 25 CAKQFNKKVR 36

## RESULT 5

TRFL MOUSE STANDARD; PRT; 707 AA.  
 AC P08071; P70690; Q61799; Q922P2;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=87280033; PubMed=3611056;  
 RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse  
 uterine secretions.";  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Morilishi K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21880-21885(1991).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; J03298; AAA40525.1; -;  
 DR EMBL; D88510; BAA13633.1; -;  
 DR EMBL; BC006904; AAH06904.1; -;  
 DR EMBL; M74778; AAA39427.1; -;  
 DR PIR; A28438; A28438.  
 DR HSP; P02788; ICB6.  
 DR MGD; MGI:96837; Ltf.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.  
 FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 505 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139 ANION (POTENTIAL).  
 FT BINDING 481 481 ANION (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2 MR -> IQG (IN REF. 1).  
 FT CONFLICT 25 25 R -> Q (IN REF. 2).  
 FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77965 MW; F26AE0340A4C19A8 CRC64;

Query Match 59.2%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 7.3;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11  
 DB 37 CLRQWQRMKKV 47





FT	METAL	111	111	IRON 1.
FT	METAL	211	211	IRON 1.
FT	METAL	272	272	IRON 1.
FT	METAL	414	414	IRON 2.
FT	METAL	452	452	IRON 2.
FT	METAL	545	545	IRON 2.
FT	METAL	614	614	IRON 2.
FT	BINDING	140	140	ANION (POTENTIAL).
FT	BINDING	482	482	ANION (POTENTIAL).
FT	CARBOHYD	252	252	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	300	300	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	495	495	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	564	564	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	708	708	AA; 08D2600AAB2F9ACD CRC64;
FT	SEQUENCE	77729	77729	NW; 08D2600AAB2F9ACD CRC64;

Query Match 53.5%; Score 38; DB 1; Length 708;  
 Best Local Similarity 54.5%; Pred. No. 36;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY	1	CFQWRNMKKV	11
DB	38	CHRWQWKKL	48

RESULT 11  
 O83A DROME STANDARD; PRT; 453 AA.  
 AC Q9VNH3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative odorant receptor 83a.  
 GN OR83A OR CGI0612.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=Berkeley;  
 MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolehakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RN CONCEPTUAL TRANSLATION.  
 RP Robergeon H.M.;  
 RA Unpublished observations (May-2001).  
 RL -!- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT  
 CC RECEPTOR.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED  
 CC RECEPTORS.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous  
 CC gene model prediction.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB03603; AAF52033.1; ALT\_SEQ.  
 DR FlyBase; FBgn0037322; Or83a.  
 DR InterPro; IPR004117; 7tm\_6.  
 DR Pfam; PF02949; 7tm\_6; 1.  
 DR Hypothetical protein; Transmembrane; G-protein coupled receptor;  
 KW Glycoprotein; Olfaction; Multigene family.  
 KW  
 FT DOMAIN 1 28  
 FT TRANSMEM 29 49  
 FT DOMAIN 50 85  
 FT TRANSMEM 86 106  
 FT DOMAIN 107 148  
 FT TRANSMEM 149 169  
 FT DOMAIN 170 203  
 FT TRANSMEM 204 224  
 FT DOMAIN 225 322  
 FT TRANSMEM 323 343  
 FT DOMAIN 344 359  
 FT TRANSMEM 360 380  
 FT DOMAIN 381 408  
 FT TRANSMEM 409 429  
 FT DOMAIN 430 453  
 FT CARBOHYD 249 249  
 SQ SEQUENCE 453 AA; 52272 MW; 4B660B3380901192 CRC64;  
 Query Match 52.1%; Score 37; DB 1; Length 453;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WORNKKVR 12  
 DB 395 WQRLKQVR 403  
 RESULT 12  
 ID VP4\_ROTGA  
 AC Q04916;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Outer capsid protein VP4 (Hemagglutinin) (Outer layer protein VP4).  
 DE S4.  
 GN S4.  
 OS Rotavirus (group B / strain ADVR) (Adult diarrhea rotavirus).  
 OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.  
 OX NCBI\_TaxID=12705;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=93233240; PubMed=8386274;  
 RA Mackow E.R., Werner-Eckert R., Fay M.E., Tao H., Chen G.-M.;  
 RT "Identification and baculovirus expression of the VP4 protein of the  
 RL human group B rotavirus ADVR.";  
 RL J. Virol. 67:2730-2738 (1993).  
 CC -!- SUBCELLULAR LOCATION: Outer capsid.  
 CC -!- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M91434; AAA47338.1; -.  
 DR InterPro; IPR000416; Cap\_VP4.  
 DR Pfam; PF00426; VP4; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 29 29  
 FT CARBOHYD 53 53  
 FT CARBOHYD 109 109  
 FT CARBOHYD 133 133  
 FT CARBOHYD 407 407  
 FT CARBOHYD 527 527  
 FT CARBOHYD 568 568  
 FT CARBOHYD 620 620  
 FT CARBOHYD 681 681  
 FT CARBOHYD 698 698  
 SQ SEQUENCE 749 AA; 84362 MW; D1223527DEAE0F21 CRC64;  
 Query Match 52.1%; Score 37; DB 1; Length 749;  
 Best Local Similarity 50.0%; Pred. No. 58;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CFQWQNNKKVR 12  
 DB 195 CFTWDMNCANVR 206  
 RESULT 13  
 ID FTSA\_HABIN  
 AC P45068;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cell division protein ftsA.  
 DE FtsA OR Hll42.  
 GN Haemophilus influenzae.  
 OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Dougherty B.A., Merrick J.M.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Tomb J.-F., Gocayne J.D.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Kiley J.M.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*  
 RT Rd.";  
 RL Science 269:496-512 (1995).  
 CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN ANOMALOUS FILAMENT  
 CC GROWTH (BY SIMILARITY). MAY BE A COMPONENT OF THE SEPTUM. IT

CC MAY INTERACT WITH FTSZ (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE FTSZ/MREB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U32794; AAC22797.1; -.  
 CC TIGR; H1142; -.  
 CC InterPro; IPR003494; FtsA.  
 CC Pfam; PF02491; FtsA; 2.  
 CC TIGRFAMs; TIGR01174; ftsA; 1.  
 CC Cell division; Cell shape; Complete proteome.  
 CC KW SEQUENCE 425 AA; 45836 MW; AF5C4B08D73CE9D CRC64;  
 CC -----  
 CC Query Match 51.4%; Score 36.5; DB 1; Length 425;  
 CC Best Local Similarity 63.6%; Pred. No. 40;  
 CC Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
 CC -----  
 CC QY 1 CPQ-WQNNKK 10  
 CC | | | | | | | | | |  
 CC Db 163 CHQDQNNKK 173  
 CC -----  
 CC RESULT 14  
 CC VNS2 DSDNV  
 CC ID VNS2 DSDNV STANDARD; PRT; 275 AA.  
 CC AC 071154;  
 CC DT 15-JUL-1999 (Rel. 38, Created)  
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 CC DE Nonstructural protein NS2.  
 CC GN NS2.  
 CC OS Diatraea saccharalis densovirus (DSDNV).  
 CC OC Viruses; ssDNA viruses; Parvoviridae; Densovirinae; Densovirus.  
 CC OX NCBI\_TaxID=72003;  
 CC RN NCBI\_TaxID=72003;  
 CC RP SEQUENCE FROM N.A.  
 CC RT Boublik Y., Kouassi K.N., Cavallaro C., Bergoin M.;  
 CC "Complete nucleotide sequence and genome organization of an infectious  
 CC clone of Diatraea saccharalis densovirus (DSDNV).";  
 CC Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF036333; AAC18000.1; -.  
 CC Nonstructural protein.  
 CC KW SEQUENCE 275 AA; 31154 MW; AA9B1B03D7718C71 CRC64;  
 CC -----  
 CC Query Match 50.7%; Score 36; DB 1; Length 275;  
 CC Best Local Similarity 85.7%; Pred. No. 32;  
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC -----  
 CC QY 4 WQNNKK 10  
 CC | | | | | | | | | |  
 CC Db 66 WQNNKK 72  
 CC -----  
 CC RESULT 15  
 CC NLA DROME  
 CC ID NLA DROME STANDARD; PRT; 292 AA.  
 CC AC Q9XZL8; Q9V391;  
 CC DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 DE NLA OR C60072.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP McCormick A.V., Goldberg M.L.;  
 RA "Gene required for elongation of meiosis I spindle in Drosophila  
 RA females";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champs M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Perriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez C.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.A., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RL "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -!- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF147700; AAD33987.1; -.  
 CC EMBL; AE003712; AAF5285.1; -.  
 CC FlyBase; FBgn026629; nla.  
 CC SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6A6CF9 CRC64;  
 CC -----

Query Match 50.7%; Score 36; DB 1; Length 292;  
Best Local Similarity 45.5%; Pred. No. 34;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQRNMKKVR 12  
||| |::|  
Db 150 FQWLRSPRRLR 160

Search completed: February 21, 2003, 07:27:58  
Job time : 5.6 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107B-83

Perfect score: 71

Sequence: 1 CFQWQRMKKV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	88.7	711	4 Q8TCD2	Q8tcd2 homo sapien
2	59	83.1	38	4 Q9UCY5	Q9ucy5 homo sapien
3	48	67.6	33	6 Q9TR80	Q9tr80 ovis aries
4	42	59.2	105	10 Q9XFD5	Q9xids oryza sativ
5	42	59.2	511	16 Q8Z462	Q8z462 salmonella
6	41	57.7	282	16 Q98Q19	Q98q19 mycoplasma
7	41	57.7	341	11 Q8R2A4	Q8r2a4 mus musculus
8	40	56.3	469	9 Q38115	Q38115 bacterioph
9	40	56.3	514	10 Q9SP27	Q9sp27 callistephu
10	40	56.3	584	17 Q9HPA3	Q9hpa3 halobacteri
11	39	54.9	205	8 Q98RR2	Q98rr2 guillardia
12	39	54.9	238	16 Q8YP77	Q8yp77 anabaena sp
13	39	54.9	332	11 Q9DAZ8	Q9daz8 mus musculus
14	39	54.9	332	11 Q91ZD5	Q91zd5 mus musculus
15	39	54.9	333	11 Q91Z75	Q91z75 mus musculus
16	39	54.9	372	10 Q81653	Q81653 hemerocalli

17	39	54.9	373	16 Q9RJP2	Q9rip2 streptomyce
18	39	54.9	531	10 Q9SICO	Q9sico arabidopsis
19	39	54.9	553	4 Q9NZL7	Q9nzl7 homo sapien
20	39	54.9	554	4 Q9NY67	Q9ny67 homo sapien
21	39	54.9	558	4 Q9UFG6	Q9ufk6 homo sapien
22	39	54.9	589	16 Q8RGT4	Q8rgt4 fusobacteri
23	39	54.9	2186	5 Q9N906	Q9n906 trypanosoma
24	38	53.5	91	15 Q77855	Q77855 human immun
25	38	53.5	91	15 Q77856	Q77856 human immun
26	38	53.5	119	15 Q8Q454	Q8q454 human immun
27	38	53.5	148	10 Q9XHP1	Q9xhpl sesamum ind
28	38	53.5	275	5 Q93780	Q93780 caenorhabdi
29	38	53.5	323	16 Q9KSU3	Q9ksu3 vibrio chol
30	38	53.5	335	5 Q61888	Q61888 caenorhabdi
31	38	53.5	393	16 Q9ZKP4	Q9zkp4 helicobacte
32	38	53.5	515	10 Q22185	Q22185 arabidopsis
33	38	53.5	543	10 Q22188	Q22188 arabidopsis
34	38	53.5	550	11 Q9UJZ5	Q9jjz5 mus musculu
35	38	53.5	932	5 Q19153	Q19153 caenorhabdi
36	38	53.5	1436	16 Q9CDI2	Q9cdi2 lactococcus
37	37	52.1	99	16 Q92GL6	Q92gl6 rickettsia
38	37	52.1	122	11 Q9D4T5	Q9d4t5 mus musculu
39	37	52.1	122	15 Q72904	Q72904 human immun
40	37	52.1	248	2 Q8RMB8	Q8rmb8 cytophaga j
41	37	52.1	270	17 Q8TNF5	Q8tnf5 methanosarc
42	37	52.1	279	16 Q8XSE2	Q8xse2 ralistonia s
43	37	52.1	283	16 Q9WYQ1	Q9wyq1 thermotoga
44	37	52.1	333	11 Q9JL96	Q9jl96 mus musculu
45	37	52.1	333	11 Q91ZP3	Q91zfp3 mus musculu

#### ALIGNMENTS

##### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.

AC Q8TCD2;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH2347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAP CRC64;

Query Match 88.7%; Score 63; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0033;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11

DB 39 CFQWQRMKKV 49

##### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.

AC Q9UCY5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;



OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 seminal plasma.";  
 RL Jpn. J. Legal Med. 49:281-293 (1995).  
 DR HSP; P02788; 18KA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F4905EBDD5B CRC64;

Query Match 83.1%; Score 59; DB 4; Length 38;  
 Best Local Similarity 90.9%; Pred. No. 0.00082;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FQWQNMKKVR 12  
 Db 21 FQWQNMKKVR 31  
 |||||:|  
 |||||:|

RESULT 3  
 Q9TR80 PRELIMINARY; PRT; 33 AA.  
 ID Q9TR80  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:23-32 (1995).  
 DR HSP; O77698; 1CE2.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAB15A73961 CRC64;

Query Match 67.6%; Score 48; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.074;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FQWQNMKKV 11  
 Db 19 CYQWQKMKRL 29  
 |||||:|  
 |||||:|

RESULT 4  
 Q9XFD5 PRELIMINARY; PRT; 105 AA.  
 ID Q9XFD5  
 AC Q9XFD5;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Cytochrome P450 (Fragment).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocarpaceae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PANICLE;  
 RA Liu J., Yang J.;  
 RT "Suppression subtractive hybridization (SSH) identified candidate

RT genes that are differentially expressed at rice young panicle.";  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 CC 1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF140486; AAD29699.1; -;  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PRO0385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 FT NON\_TER  
 SQ SEQUENCE 105 AA; 11912 MW; B0EEFCD487E19F9 CRC64;

Query Match 59.2%; Score 42; DB 10; Length 105;  
 Best Local Similarity 70.0%; Pred. No. 3.2;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWQNMKK 10  
 Db 61 CFQWQNLGKK 70  
 |||||:|  
 |||||:|

RESULT 5  
 Q8Z462 PRELIMINARY; PRT; 511 AA.  
 ID Q8Z462  
 AC Q8Z462;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein STY3070.  
 GN STY3070.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 Feltham T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 Whitehead S., Barrall B.G.;  
 RA "Complete genome sequence of a multiple drug resistant Salmonella  
 enterica serovar Typhi CT18.";  
 RT Nature 413:848-852 (2001).  
 DR EMBL; AL627276; CAD06049.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 511 AA; 58126 MW; E2DD124E10D1788 CRC64;

Query Match 59.2%; Score 42; DB 16; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CFQWQNMKKVR 12  
 Db 350 CFQWQNMKKVR 361  
 |||||:|  
 |||||:|

RESULT 6  
 Q98Q19 PRELIMINARY; PRT; 282 AA.  
 ID Q98Q19  
 AC Q98Q19;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE tRNA pseudouridine synthase B (tRNA pseudouridine 55 synthase) (PSI55  
 synthase) (Pseudouridylylase synthase) (Uracil hydrolyase)  
 DE (EC 4.2.1.70).  
 GN MYPV\_5500.

```

OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dywig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445565; CAC13723.1; -.
DR Mypulist; MIPU_5500; -.
DR InterPro; IPR004510; TruB.
DR InterPro; IPR002501; TruB N.
DR Pfam; PF01509; TruB N; 1.
DR TIGRPFAMs; TIGR00431; TruB; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 282 AA; 31961 MW; A598529F52B1EBEE CRC64;

Query Match 57.7%; Score 41; DB 16; Length 282;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQRNMKKV 11
| | | | |
DB 20 FAWQNNIKKI 29

RESULT 7
Q8R2A4 PRELIMINARY; PRT; 341 AA.
AC Q8R2A4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vomeronasal receptor VIR56.
GN VIR56.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129X1/SVJ;
RX MEDLINE=21676859; PubMed=11802169;
RA Rodriguez I., Del Punta K., Rothman A., Ishii T., Mombaerts P.;
RT "Multiple new and isolated families within the mouse superfamily of
RT V1r vomeronasal receptors.";
RL Nat. Neurosci. 5:134-140(2002).
DR EMBL; AY065506; AAL47911.1; -.
KW Receptor.
SQ SEQUENCE 341 AA; 39632 MW; B3D8259F7AD11130 CRC64;

Query Match 57.7%; Score 41; DB 11; Length 341;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRNMKK 9
| | | | |
DB 331 CFTWTRNIK 339

RESULT 8
Q38115 PRELIMINARY; PRT; 469 AA.
AC Q38115;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 08, Last annotation update)

```

```

DE ORF29.
OS Bacteriophage rlt.
OC Viruses.
OX NCBI_TaxID=43685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332668; PubMed=8730874;
RA Nauta A., Van Sinderen D., Karsens H., Smit E., Venema G., Kok J.;
RT "Inducible gene expression mediated by a repressor-operator system
RT isolated from Lactococcus lactis bacteriophage rlt.";
RL Mol. Microbiol. 19:1331-1341(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332669; PubMed=8730875;
RA Van Sinderen D., Karsens H., Kok J., Terpstra P., Ruiters M.H.,
RA Venema G., Nauta A.;
RT "Sequence analysis and molecular characterization of the temperate
RT lactococcal bacteriophage rlt.";
RL Mol. Microbiol. 19:1343-1355(1996).
DR EMBL; U38906; AAB18704.1; -.
SQ SEQUENCE 469 AA; 53160 MW; 1F8E02D4325C6BB9 CRC64;

Query Match 56.3%; Score 40; DB 9; Length 469;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRNMKK 10
| | | | |
DB 39 CYPWQKLLK 48

RESULT 9
Q9SP27 PRELIMINARY; PRT; 514 AA.
AC Q9SP27;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Flavone synthase II.
GN CYP93B5.
OS Callistephus chinensis (China aster).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Asteraceae; Callistephus.
OX NCBI_TaxID=13379;
RN [1]
RP SEQUENCE FROM N.A.
RA Martens S., Forkmann G.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF198612; AAF04115.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 514 AA; 58412 MW; 90B631B28952A5E7 CRC64;

Query Match 56.3%; Score 40; DB 10; Length 514;
Best Local Similarity 54.5%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRNMKKV 11
| | | | |
DB 470 CFEMWANDKEV 480

RESULT 10
Q9HPA3 PRELIMINARY; PRT; 584 AA.
AC Q9HPA3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE VNG1732C.  
 GN VNG1732C.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=11016950;  
 RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
 RA Shukla H.D., Laasy S.R., Baliga N.S., Thorason V., Shrogha J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,  
 RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W., Dale H.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M.,  
 RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AE005078; AAC19967.1; -;  
 DR InterPro; IPR001646; 5peptide repeat.  
 DR InterPro; IPR001622; K-channel pore.  
 DR Pfam; PF00805; Pentapeptide; 2.  
 KW Complete proteome.  
 SQ SEQUENCE 584 AA; 65151 MW; 21BF5D5F0486CCC6 CRC64;

Query Match 56.3%; Score 40; DB 17; Length 584;  
 Best Local Similarity 41.7%; Pred. No. 45;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQQRNMKKVR 12  
 |||:::|:  
 Db 445 CFTWRKDMERK 456

RESULT 11  
 Q98R2 PRELIMINARY; PRT; 205 AA.  
 ID Q98R2  
 AC Q98R2;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE 265 proteasome SU B5.  
 DE PRS5.  
 GN Guillardia theta (Cryptomonas phi).  
 OS Nucleomorph.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21223349; PubMed=11323671;  
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,  
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;  
 RT "The highly reduced genome of an enslaved algal nucleus."  
 RL Nature 410:1091-1096(2001).  
 DR EMBL; AF165818; AAK39885.1; -;  
 DR InterPro; IPR000243; Proteasome B.  
 DR InterPro; IPR001353; Proteasome protease.  
 DR Pfam; PF00227; proteasome; 1.  
 DR PRINTS; PR00141; PROTEASOME.  
 KW Proteasome.  
 SQ SEQUENCE 205 AA; 22691 MW; D30F5289C8C85049 CRC64;

Query Match 54.9%; Score 39; DB 8; Length 205;  
 Best Local Similarity 45.5%; Pred. No. 23;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQQRNMKKV 11  
 |||:::|:  
 Db 63 CFFWRNLSL 73

RESULT 12  
 Q8YP77 PRELIMINARY; PRT; 298 AA.  
 ID Q8YP77  
 AC Q8YP77;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein ALR4323.  
 GN ALR4323.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120."  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003596; BAB76022.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 298 AA; 34513 MW; 64036E6B52299A9F CRC64;

Query Match 54.9%; Score 39; DB 16; Length 298;  
 Best Local Similarity 66.7%; Pred. No. 34;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FQWRNMKK 10  
 |||:::|:  
 Db 163 FHWQRNVRK 171

RESULT 13  
 Q9DAZ8 PRELIMINARY; PRT; 332 AA.  
 ID Q9DAZ8  
 AC Q9DAZ8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE 160000123Rik protein.  
 DE 160000123Rik.  
 GN Mus musculus (Mouse).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CS7BL/6J; TISSUE=PLACENTA;  
 RC MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Sakota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";

```

RL Nature 409:685-690(2001).
DR EMBL; AK005389; BAB23995.1; -.
DR HSP; P07711; 1CJL.
DR MEROPS; C01.053; -.
DR MGD; MGI:1916256; 1600000I123Rik.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00112; Peptidase C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR ProDom; PD000158; Peptidase C1; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN 1.
SQ SEQUENCE 332 AA; 37298 MW; 0804F1EA5B653BE0 CRC64;

Query Match          54.9%; Score 39; DB 11; Length 332;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQNMKKVR 12
DB 52 WEENMKKIK 60

RESULT 14
Q91ZDS PRELIMINARY; PRT; 332 AA.
AC Q91ZDS;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cathepsin-3 precursor.
GN CTS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RA Deussing J., Kouadio M., Rehman S., Werber I., Schwinde A., Peters C.;
RT "Identification and Characterization of a Dense Cluster of Placenta-
RT specific Cysteine Peptidases and Related Genes on Mouse Chromosome
RT 13.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034574; AAK58450.1; -.
DR MGD; MGI:2151929; Cts3.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00112; Peptidase C1; 1.
DR ProDom; PD000158; Peptidase C1; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN 1.
KW Signal.
FT SIGNAL
FT CHAIN
FT CHAIN 114 332 CATHEPSIN-3.
SQ SEQUENCE 332 AA; 37326 MW; 4184B90725B41C0D CRC64;

Query Match          54.9%; Score 39; DB 11; Length 332;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQNMKKVR 12
DB 52 WEENMKKIK 60

RESULT 15
Q91Z75 PRELIMINARY; PRT; 333 AA.
AC Q91Z75;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cathepsin M.
GN CTSM.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVEVTAcfBR; TISSUE=SPLLEN;
RA Rehman S., Peters C., Deussing J.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057446; RAL15416.1; -.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00112; Peptidase C1; 1.
DR ProDom; PD000158; Peptidase C1; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN 1.
SQ SEQUENCE 333 AA; 37388 MW; 6DD0BEB91C033110 CRC64;

Query Match          54.9%; Score 39; DB 11; Length 333;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQNMKKVR 12
DB 52 WEENMKKIK 60

Search completed: February 21, 2003, 07:44:37
Job time : 21.8 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 Seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107B-84  
Perfect score: 71  
Sequence: 1 CFQWERNMKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq 101002.\*  
1: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SID22/cgcdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SID22/cgcdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SID22/cgcdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SID22/cgcdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	100.0	12	AA1980	Human lactoferrin
2	68	95.8	12	AA1981	Human lactoferrin
3	68	95.8	12	AA1982	Human lactoferrin
4	68	95.8	12	AA1983	Human lactoferrin
5	68	95.8	13	AA1984	Human lactoferrin
6	68	95.8	13	AA1985	Human lactoferrin
7	68	95.8	13	AA1986	Human lactoferrin
8	68	95.8	14	AA1987	Human lactoferrin
9	68	95.8	14	AA1988	Human lactoferrin
10	68	95.8	14	AA1989	Human lactoferrin

11	68	95.8	15	17	AA198554	Peptide for anti-u
12	68	95.8	15	21	AA198035	Human lactoferrin
13	68	95.8	15	21	AA198062	Human lactoferrin
14	68	95.8	15	21	AA198063	Human lactoferrin
15	68	95.8	16	21	AA198031	Human lactoferrin
16	68	95.8	16	21	AA198064	Human lactoferrin
17	68	95.8	16	21	AA198065	Human lactoferrin
18	68	95.8	17	21	AA198034	Human lactoferrin
19	68	95.8	17	21	AA198066	Human lactoferrin
20	68	95.8	17	21	AA198067	Human lactoferrin
21	68	95.8	18	15	AA198352	Human lactoferrin
22	68	95.8	18	17	AA198397	Advanced glycosyla
23	68	95.8	18	21	AA198033	Human lactoferrin
24	68	95.8	19	21	AA198867	Amino acid sequenc
25	68	95.8	19	21	AA198032	Human lactoferrin
26	68	95.8	20	13	AA198180	Anti microbial pep
27	68	95.8	20	14	AA198484	Lactoferrin-relate
28	68	95.8	20	15	AA198530	Lactoferrin derive
29	68	95.8	20	15	AA198531	Lactoferrin derive
30	68	95.8	20	15	AA198541	Lactoferrin derive
31	68	95.8	20	15	AA198542	Lactoferrin derive
32	68	95.8	20	16	AA198468	Bovine lactoferrin
33	68	95.8	20	16	AA198469	Bovine lactoferrin
34	68	95.8	20	16	AA1980263	Anti-parasitic lac
35	68	95.8	20	16	AA1980264	Anti-parasitic lac
36	68	95.8	20	17	AA198553	Peptide for anti-u
37	68	95.8	20	17	AA198185	Lactoferrin-derive
38	68	95.8	20	17	AA1980345	Lactoferrin-derive
39	68	95.8	20	17	AA1980607	Lactoferrin derive
40	68	95.8	20	17	AA1987621	Lactoferrin-derive
41	68	95.8	20	17	AA1987622	Lactoferrin-derive
42	68	95.8	20	18	AA1986150	Lactoferrin derive
43	68	95.8	20	18	AA1986150	Lactoferrin derive
44	68	95.8	20	18	AA1984036	Anti-parasitic pep
45	68	95.8	20	19	AA1980310	Thrombus formation
						Lactoferrin hydrol

ALIGNMENTS

RESULT 1  
AA198084  
ID AA198084 standard; Peptide; 12 AA.  
AC AA198084;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:84.

Human; lactoferrin; modification; infection; inflammation; tumour;  
food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
urinary tract infection; colitis; Candida infection; fungicidal;  
bactericidal; preservative.

XX Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.

PD 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

PA (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 22; Page 36; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 100.0%; Score 71; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05; Indels 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWERNRKVR 12  
 Db 1 CFQWERNRKVR 12  
 |||||:|||||  
 RESULT 2  
 AAY78038  
 ID AAY78038 standard; Peptide; 12 AA.  
 XX  
 AC AAY78038;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:38.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 70; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWERNRKVR 12  
 Db 1 CFQWERNRKVR 12  
 |||||:|||||  
 RESULT 3  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 XX  
 AC AAY78046;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:46.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 15; Page 35; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMRKVR 12  
 ||||:|||||  
 Db 1 CFQWERNMRKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:47.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX PS WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.3e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMRKVR 12  
 ||||:|||||  
 Db 1 CFQWERNMRKVR 12

RESULT 5  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:37.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNRKVR 12  
 Db ||||:|||||  
 2 CFQWQNRNKRVR 13

## RESULT 6

AAV78048  
 ID AAY78048 standard; Peptide; 13 AA.

XX AAY78048;  
 AC  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:48.

DE Human, lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

QY Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNRKVR 12  
 Db ||||:|||||  
 2 CFQWQNRNKRVR 13

## RESULT 7

AAV78049

ID AAY78049 standard; Peptide; 13 AA.

XX AAY78049;

AC

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

DE Human, lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

QY Query Match 95.8%; Score 68; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNRKVR 12  
 Db ||||:|||||  
 2 CFQWQNRNKRVR 13

## RESULT 8

AAV78036

ID AAY78036 standard; Peptide; 14 AA.

XX AAY78036;

AC

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.



XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147388/13.  
 XX CC New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 69; 102pp; English.  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 5.1e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWERNMKVR 12  
 Db ||||:|||||  
 3 CFQWERNMKVR 14  
 RESULT 9  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AC AAY78050;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:50.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147388/13.  
 XX CC New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 15; Page 75; 102pp; English.  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 95.8%; Score 68; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 5.1e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWERNMKVR 12  
 Db ||||:|||||  
 3 CFQWERNMKVR 14  
 RESULT 10  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX AC AAY78051;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:51.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.



Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 5.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMRKVR 12  
 ||||:|||||  
 Db 4 CFQWRNMRKVR 15

## RESULT 13

AAV78062  
 ID AAV78062 standard; Peptide; 15 AA.

XX AAV78062;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:62.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

DR WPI; 2000-147388/13.

PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 15; Page 81; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 5.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMRKVR 12  
 ||||:|||||  
 Db 4 CFQWRNMRKVR 15

## RESULT 14

AAV78063  
 ID AAV78063 standard; Peptide; 15 AA.

XX AAV78063;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:63.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

PD 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

DR WPI; 2000-147388/13.

PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 18; Page 81; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 95.8%; Score 68; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 5.5e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMRKVR 12  
 ||||:|||||  
 Db 4 CFQWRNMRKVR 15

## RESULT 15

AAV78031  
 ID AAV78031 standard; Peptide; 16 AA.

XX AAV78031;

XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:31.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX

OS Homo sapiens.  
 OS Synthetic.  
 XX

FN WO200001730-A1.  
 XX

PD 13-JAN-2000.  
 XX

PF 06-JUL-1999; 99WO-SE01230.  
 XX

PR 06-JUL-1999; 98SE-0002441.  
 PR

17-JUL-1998; 98SE-0002562.  
 PR

29-DEC-1998; 98SE-0004614.  
 XX

PA (ASCI-) A+ SCI INVEST AB.  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX

DR WPI; 2000-147388/13.  
 XX

PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PS Claim 11; Page 68; 102pp; English.  
 XX

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 CC

XX Sequence 16 AA;  
 SQ

Query Match 95.8%; Score 68; DB 21; Length 16;  
 Best Local Similarity 91.7%; Pred. No. 5.9e-05;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNRKVR 12  
 |||||

Db 5 CFQWQNNRKVR 16  
 |||||

Search completed: February 21, 2003, 07:37:15  
 Job time : 29.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 Seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-84

Perfect score: 71

Sequence: 1 CFQERNRMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA\*

1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	18	1	US-08-204-487-3
2	68	95.8	18	2	US-08-485-948-8
3	68	95.8	18	2	US-08-628-380-8
4	68	95.8	18	2	US-08-475-055-8
5	68	95.8	20	1	US-07-755-161A-3
6	68	95.8	20	1	US-07-891-174-3
7	68	95.8	20	1	US-08-204-487-1
8	68	95.8	20	1	US-08-256-771-24
9	68	95.8	20	1	US-08-256-771-25
10	68	95.8	20	1	US-08-381-984-24
11	68	95.8	20	1	US-08-381-984-25
12	68	95.8	22	4	US-09-508-734-4
13	68	95.8	24	4	US-09-508-734-6
14	68	95.8	25	1	US-07-755-161A-10
15	68	95.8	25	1	US-07-891-174-10
16	68	95.8	25	1	US-08-204-487-7
17	68	95.8	29	4	US-09-508-734-8
18	68	95.8	36	1	US-07-755-161A-8
19	68	95.8	36	1	US-07-891-174-8
20	68	95.8	36	1	US-08-256-771-30
21	68	95.8	36	1	US-08-381-984-29
22	68	95.8	47	2	US-08-464-182A-6
23	68	95.8	47	2	US-08-406-271-6
24	68	95.8	50	2	US-08-693-274A-7
25	68	95.8	52	4	US-09-017-043A-3
26	68	95.8	53	2	US-08-464-182A-5
27	68	95.8	53	2	US-08-406-271-5

28	68	95.8	54	2	US-08-464-182A-2
29	68	95.8	54	2	US-08-406-271-2
30	68	95.8	694	3	US-08-724-586-2
31	68	95.8	694	4	US-09-421-632-2
32	68	95.8	694	4	US-09-932-190-2
33	68	95.8	705	2	US-08-655-640-2
34	68	95.8	708	2	US-08-655-640-4
35	68	95.8	711	1	US-08-154-019-4
36	68	95.8	711	1	US-08-461-333-4
37	68	95.8	711	3	US-08-464-167-4
38	68	95.8	711	3	US-09-158-313-4
39	68	95.8	711	4	US-08-476-798-4
40	65	91.5	711	1	US-08-145-681-2
41	65	91.5	711	1	US-08-250-308-2
42	65	91.5	711	1	US-08-453-703-2
43	65	91.5	711	2	US-08-456-106-2
44	65	91.5	711	3	US-08-456-108-2
45	65	91.5	711	4	US-09-265-577-2

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 95.8%; Score 68; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
||||:|||||  
DB 1 CFQWQNRMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VIASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/485,948  
FILING DATE: JUNE 7, 1995  
CLASSIFICATION: 436

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995

APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995

## CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-C1, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

## US-08-485-948-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
||||:|||||  
DB 1 CFQWQNRMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341

## GENERAL INFORMATION:

APPLICANT: LI, YONG MING  
APPLICANT: VIASSARA, HELEN  
APPLICANT: CERAMI, ANTHONY  
TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
NUMBER OF SEQUENCES: 9

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,380  
FILING DATE: April 4, 1996  
CLASSIFICATION: 514

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 514

## ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-C1, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

## US-08-628-380-8

Query Match 95.8%; Score 68; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
||||:|||||

DB 1 CFQWQNRMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VIASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/488,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LF-C1, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

```

```

Query Match          95.8%; Score 68; DB 2; Length 18;
Best Local Similarity 91.7%; Pred No. 4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CFQWRNNRKVR 12
    |||:|||||
Db 1 CFQWRNNRKVR 12

```

## RESULT 5

```

US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633

```

```

GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible

```

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3

```

Query Match

95.8%; Score 68; DB 1; Length 20;

```
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12
Db 2 CFQWQNNMKVR 13

RESULT 6
US-07-891-174-3
; Sequence 3, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site

; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-3

Query Match 95.8%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12
Db 2 CFQWQNNMKVR 13

RESULT 7
US-08-204-487-1
; Sequence 1, Application US/08204487
; Patent No. 5565425
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, NAOKI
; APPLICANT: NAKASHIMA, HIDEKI
; APPLICANT: MOSUCHI, WATARU
; APPLICANT: TANAKA, SHIGERU
; APPLICANT: DOSAKO, SHUN-ICHI
; APPLICANT: KAWASAKI, YOSHIHIRO
; APPLICANT: UCHIDA, TOSHIKI
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,487
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FUN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
```



```

; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
; OTHER INFORMATION: DERIVED FROM HUMAN LACTOFERRIN"
; US-08-204-487-1

Query Match 95.8%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||:|||||
Db 2 CFQWRNMRKVR 13

RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"

```

```

US-08-256-771-24

Query Match 95.8%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||:|||||
Db 2 CFQWRNMRKVR 13

RESULT 9
US-08-256-771-25
; Sequence 25, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are protected to
; OTHER INFORMATION: prevent disulfide bond"
; US-08-256-771-25

Query Match 95.8%; Score 68; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||:|||||
Db 2 CFQWRNMRKVR 13

RESULT 10
US-08-381-984-24

```

```
; Sequence 24, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof
; US-08-381-984-24
;
; Query Match 95.8%; Score 68; DB 1; Length 20;
; Best Local Similarity 91.7%; Pred. No. 4.4e-05;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CFQWERNMRKVR 12
; Db 2 CFQWQNRMRKVR 13
;
; RESULT 11
; US-08-381-984-25
; Sequence 25, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
```

```
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof
; US-08-381-984-25
;
; Query Match 95.8%; Score 68; DB 1; Length 20;
; Best Local Similarity 91.7%; Pred. No. 4.4e-05;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CFQWERNMRKVR 12
; Db 2 CFQWQNRMRKVR 13
;
; RESULT 12
; US-09-508-734-4
; Sequence 4, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; TITLE OF INVENTION: useful microorganism thereof
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KE99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 4
```

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match          95.8%; Score 68; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 4.9e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          95.8%; Score 68; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 5.3e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 3 CFQWRNMRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; NUMBER OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500KB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note="thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note="thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match          95.8%; Score 68; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 5.5e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 4 CFQWRNMRKVR 15
```

RESULT 15  
US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLER:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10  
; Query Match 95.8%; Score 68; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 5.5e-05;  
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
; Qy 1 CFQWRNNRKVR 12  
; |||:|||||  
; Db 4 CFQWRNNRKVR 15  
; Search completed: February 21, 2003, 07:50:36  
; Job time : 8.7 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-84

Perfect score: 71 CFQWERNMKVR 12

Sequence: 1 CFQWERNMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT NEW PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07 PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/ECTUS PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09 PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10 PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	15	9	US-09-798-869-2
2	68	95.8	25	9	US-09-798-869-20
3	68	95.8	694	9	US-10-023-096-2
4	60	84.5	15	9	US-09-798-869-6
5	51	71.8	15	9	US-09-798-869-3
6	51	71.8	25	9	US-09-798-869-23
7	43	60.6	15	9	US-09-798-869-7
8	42	59.2	15	9	US-09-798-869-4
9	42	59.2	25	9	US-09-798-869-22
10	39	54.9	15	9	US-09-798-869-8
11	39	54.9	15	9	US-09-798-869-29
12	39	54.9	15	9	US-09-798-869-30
13	38	53.5	489	9	US-09-888-320-2
14	37	52.1	846	9	US-10-051-409-4
15	36	50.7	21	10	US-09-864-761-47985
16	36	50.7	40	10	US-09-864-761-46393
17	35	49.3	86	9	US-09-738-626-5715
18	35	49.3	333	9	US-09-796-753-26
19	35	49.3	338	9	US-09-978-295A-119

#### SUMMARIES

20	35	49.3	338	9	US-09-978-697-119	Sequence 119, App
21	35	49.3	338	9	US-09-978-192A-119	Sequence 119, App
22	35	49.3	338	9	US-09-999-832A-119	Sequence 119, App
23	35	49.3	338	9	US-09-978-189-119	Sequence 119, App
24	35	49.3	388	10	US-09-989-861-8	Sequence 8, Appli
25	35	49.3	553	9	US-09-796-753-14	Sequence 14, Appli
26	35	49.3	553	10	US-09-981-649A-6	Sequence 6, Appli
27	35	49.3	553	10	US-09-981-649A-24	Sequence 24, Appli
28	35	49.3	554	10	US-09-981-649A-30	Sequence 30, Appli
29	35	49.3	554	10	US-09-981-649A-32	Sequence 32, Appli
30	35	49.3	559	10	US-09-981-649A-28	Sequence 28, Appli
31	35	49.3	1701	9	US-09-963-959-2	Sequence 2, Appli
32	34	47.9	50	10	US-09-864-761-46806	Sequence 46806, A
33	34	47.9	62	10	US-09-815-242-12129	Sequence 12129, A
34	34	47.9	62	10	US-09-815-242-13026	Sequence 13026, A
35	34	47.9	63	9	US-09-984-245-320	Sequence 320, App
36	34	47.9	66	10	US-09-864-761-33619	Sequence 33619, A
37	34	47.9	248	10	US-09-925-300-1324	Sequence 1324, Ap
38	34	47.9	747	9	US-10-066-500-58	Sequence 58, Appli
39	34	47.9	747	9	US-10-002-796-58	Sequence 58, Appli
40	34	47.9	747	9	US-10-066-273-58	Sequence 58, Appli
41	34	47.9	747	9	US-10-066-494-58	Sequence 58, Appli
42	33	46.5	77	10	US-09-864-761-41002	Sequence 41002, A
43	33	46.5	88	10	US-09-764-869-1200	Sequence 1200, Ap
44	33	46.5	95	10	US-09-764-864-1031	Sequence 1031, Ap
45	33	46.5	168	10	US-09-925-302-530	Sequence 530, App

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 95.8%; Score 68; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWERNMKVR 12  
Db 3 CFQWERNMKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 95.8%; Score 68; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 5.8e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
Db 3 CFQWERNMKVR 14

## RESULT 3

US-10-023-096-2  
Sequence 2, Application US/10023096  
Patent No. US20020160941A1

GENERAL INFORMATION:  
APPLICANT: Kruzel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,096  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William B.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-10-023-096-2

Query Match 95.8%; Score 68; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0014;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
Db 22 CFQWERNMKVR 33

## RESULT 4

US-09-798-869-6  
Sequence 6, Application US/09798869  
Publication No. US20030022821A1

GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 84.5%; Score 60; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.0007;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
Db 3 CFQWERNMKVR 14

## RESULT 5

US-09-798-869-3  
Sequence 3, Application US/09798869  
Publication No. US20030022821A1

GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15  
TYPE: PRT  
ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 71.8%; Score 51; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWERNMRKV 11  
|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 71.8%; Score 51; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.032;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWERNMRKV 11  
|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 60.6%; Score 43; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.38;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWERNMRKV 11  
|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.55;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWERNMRKV 11  
|:|:|:|:|:  
Db 3 CLRQWERNMRKV 13

## RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 59.2%; Score 42; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 0.9;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWERNMRKV 11  
|:|:|:|:|:  
Db 3 CLRQWERNMRKV 13

RESULT 12  
US-09-798-869-30

RESULT 14  
US-10-051-409-4  
Sequence 4, Application US/10051409  
Publication No. US20030027171A1  
GENERAL INFORMATION.



; APPLICANT: Yang, Chu-Wen  
; APPLICANT: Tsou, Ann-Ping  
; APPLICANT: Chi, Chin-Wen  
; APPLICANT: Fann, Ming-Ji  
; APPLICANT: Chou, Chen-Kung  
; TITLE OF INVENTION: CELL CYCLE REGULATOR PROTEIN  
; FILE REFERENCE: 12005-003001  
; CURRENT APPLICATION NUMBER: US/10/051,409  
; CURRENT FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: 60/262,885  
; PRIOR FILING DATE: 2001-01-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 846  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-051-409-4

Query Match 52.1%; Score 37; DB 9; Length 846;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CFQWERNM 8  
||:|:|:  
Db 448 CFWDRL 455

RESULT 15  
US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Shaaron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL096701.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EST\_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06  
US-09-864-761-47985

Query Match 50.7%; Score 36; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 7;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CFQWER 6  
|||||  
Db 16 CFQWRR 21

Search completed: February 21, 2003, 08:08:08  
Job time : 10.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-84  
Perfect score: 71  
Sequence: 1 CFQWERNMKXVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	711	TFHUL	lactotransferrin p
2	51	71.8	708	JC2323	lactoferrin - goat
3	48	67.6	33	S52107	lactoferrin - sheep
4	44	62.0	511	A30858	hypothetical prote
5	42	59.2	707	A28438	lactoferrin precu
6	41	57.7	205	E90094	26S proteasome SU
7	41	57.7	275	T22597	hypothetical prote
8	41	57.7	1432	B85431	trichosalin like
9	41	57.7	4568	T08030	dynamin beta heavy
10	40	56.3	274	B09050	apolipoprotein B-1
11	40	56.3	289	G86403	33.3K hypothetical
12	40	56.3	361	T29571	hypothetical prote
13	39	54.9	298	A23346	hypothetical prote
14	39	54.9	365	I37477	MHC class I histoc
15	39	54.9	749	A45687	outer capsid prote
16	39	54.9	820	G82168	trimethylamine-N-o
17	38	53.5	393	D71876	3-deoxy-manno-octu
18	38	53.5	433	T32605	hypothetical prote
19	38	53.5	481	B96691	probable cytochrom
20	38	53.5	489	C70655	probable monooxyge
21	38	53.5	492	B85441	cytochrome P450-11
22	38	53.5	499	T04730	cytochrome P450 ho
23	38	53.5	500	C85441	cytochrome P450-11
24	38	53.5	500	T52174	cytochrome P450 mo
25	38	53.5	584	C84325	hypothetical prote
26	38	53.5	587	H96515	hypothetical prote
27	38	53.5	932	T28820	hypothetical prote
28	37	52.1	152	B64485	hypothetical prote
29	37	52.1	206	H97451	pyridoxamine 5'-ph

30	37	52.1	206	2	AB2670	pyridoxamine 5'-ph
31	37	52.1	208	2	AG3441	probable pyridoxam
32	37	52.1	255	2	B87515	signal peptidase I
33	37	52.1	334	1	KHRTL	cathepsin L (EC 3.
34	37	52.1	334	1	KHMSL	cathepsin L (EC 3.
35	37	52.1	415	2	C71467	probable tyrosine
36	37	52.1	505	2	T10896	cytochrome P450 (E
37	37	52.1	517	2	F71417	cytochrome P450 -
38	37	52.1	536	2	T24218	hypothetical prote
39	37	52.1	681	2	T19429	hypothetical prote
40	37	52.1	720	2	A96807	hypothetical prote
41	36	50.7	124	2	C96582	F151.22 [imported
42	36	50.7	134	2	B87162	probable secreted
43	36	50.7	303	2	E70848	probable oxidoredu
44	36	50.7	339	2	T09217	protein sam2B - sp
45	36	50.7	433	2	B90419	hypothetical prote

## ALIGNMENTS

## RESULT 1

## TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74;

R:Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R:Rev. W.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, S288, 1990

A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>

A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Yang, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A:Title: Differential molecular mechanism of the estrogen action that regulates lactoferrin

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A:Experimental source: placenta

A>Note: sequence extracted from NCBI backbone (NCBIP:122202)

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A:Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <ST1>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-28, 'X', 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:G186815; PIDN:AAA86665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SWKPVN' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Mez-Bouigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LTF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;23-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status e  
 F;157,498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat  
 Query Match 95.8%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00046;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CFQWRNRKVR 12  
 |||:|||||  
 Db 39 CFQWRNRKVR 50  
 RESULT 2  
 JC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JC2323  
 R;Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: JC2323; MUID:94380047; PMID:8093048  
 A;Accession: JC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.8%; Score 51; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.46;  
 Matches 7; Conservative 3; Mismatches 0; Gaps 0;

QY 1 CFQWRNRKVR 11  
 |||:|||||  
 Db 38 CYQWRNRKRL 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samouri, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 67.6%; Score 48; DB 2; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.067;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNRKVR 11  
 |||:|||||  
 Db 19 CYQWRNRKRL 29

RESULT 4  
 AB0858  
 hypothetical protein STY3070 [imported] - Salmonella enterica subsp. enterica serovar Ty  
 C;Species: Salmonella enterica subsp. enterica serovar Typhi  
 A;Note: this species has also been called Salmonella typhi  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 C;Accession: AB0858  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moulle, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A;Reference number: AB0502; PMID:11677608  
 A;Accession: AB0858  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-511 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:G16504016; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STY3070

Query Match 62.0%; Score 44; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 5.7;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRNRKVR 12  
 |||:|||||  
 Db 350 CFQWRNRKVR 361

RESULT 5  
 A28438  
 lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory cells  
 A;Reference number: A92596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042039; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 59.2%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWERNMKV 11  
 | : : : :  
 Db 37 CLRQWQEMK 47

RESULT 6  
 E90094  
 26S proteasome SU B5 [imported] - Guillardia theta nucleomorph  
 C;Species: nucleomorph Guillardia theta  
 A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 A;Accession: E90094  
 R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reishus, S.  
 Nature 410, 1091-1096, 2001  
 A;Title: The highly reduced genome of an enslaved algal nucleus.  
 A;Reference number: A99082; MUID:11323671; PMID:11323671  
 A;Accession: E90094  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-205 <DOU>  
 A;Cross-references: GB:AF165818; NID:gl3794510; PIDN:AAK39885.1; GSPDB:GN00150  
 C;Genetics:  
 A;Gene: prsB5  
 A;Map position: 1  
 A;Genome: nucleomorph  
 C;Keywords: nucleomorph

Query Match 57.7%; Score 41; DB 2; Length 205;  
 Best Local Similarity 75.0%; Pred. No. 7.6;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWERNM 8  
 | : : : :  
 Db 63 CFFWERNL 70

RESULT 7  
 T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C;Accession: T22597  
 R;Dobson, R.  
 submitted to the EMBL Data Library, October 1996  
 A;Reference number: Z19587

A;Accession: T22597  
 A;Status: preliminary; translated from GB/EMBL/DDBY  
 A;Molecule type: DNA  
 A;Residues: 1-275 <WIL>  
 A;Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4  
 A;Experimental source: clone F53H4  
 C;Genetics:  
 A;Gene: CESP:F53H4.4  
 A;Map position: X  
 A;Introns: 6/1; 153/1  
 C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 57.7%; Score 41; DB 2; Length 275;  
 Best Local Similarity 63.6%; Pred. No. 10;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWERNMKVR 12  
 | : : : :  
 Db 262 FQWKSMRKTR 272

RESULT 8  
 B85431  
 trichogyalin like protein [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 A;Accession: B85431  
 R;Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory  
 Nature 402, 769-777, 1999  
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A;Reference number: A85001; MUID:20083488; PMID:10617198  
 A;Accession: B85431  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1432 <STO>  
 A;Cross-references: GB:NC\_001268; NID:g7270600; PIDN:CAB80318.1; GSPDB:GN00140  
 C;Genetics:  
 A;Gene: AT4g36520  
 A;Map position: 4

Query Match 57.7%; Score 41; DB 2; Length 1432;  
 Best Local Similarity 54.5%; Pred. No. 56;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWERNMKVR 12  
 | : : : :  
 Db 583 YDWEQNARKLR 593

RESULT 9  
 T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C;Species: Chlamydomonas reinhardtii  
 C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 A;Accession: T08030  
 R;Mitchell, D.R.; Brown, K.S.  
 J. Cell Sci. 107, 635-644, 1994  
 A;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
 A;Reference number: Z16302; MUID:94274778; PMID:8006077  
 A;Accession: T08030  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-4568 <MIT>  
 A;Cross-references: EMBL:U02963; NID:g409965; PIDN:AAA19956.1; PID:G514215  
 A;Experimental source: strain 21gr  
 C;Genetics:  
 A;Gene: ODA4  
 A;Map position: IX  
 A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3; 1334/3; 3686/3; 3882/3; 4240/3  
 C;Superfamily: dynein heavy chain, ciliary  
 C;Keywords: nucleotide binding; P-loop  
 F;1919-1926/Region: nucleotide-binding motif A (P-loop)

F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 57.7%; Score 41; DB 2; Length 4568;  
Best Local Similarity 41.7%; Pred. No. 1.8e+02;  
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
|||: : : :  
Db 1852 CFQWSQLRYIQ 1863

## RESULT 10

B60950  
apolipoprotein B-100 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1993 #sequence\_revision 09-Sep-1994 #text\_change 07-Oct-1994

C:Accession: B60950

R:Law, A.; Scott, J.

J. Lipid Res. 31, 1109-1120, 1990

A:Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL

A:Reference number: A60950; MUID:90324804; PMID:2373961

A:Accession: B60950

A:Molecule type: mRNA

A:Residues: 1-274 <LAW>

A>Note: authors translated the codon ATA for residue 8 as Val

C:Superfamily: apolipoprotein B

C:Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 56.3%; Score 40; DB 2; Length 274;  
Best Local Similarity 56.7%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WERNMKVR 12  
|:|:|:|:|  
Db 39 WDRNLRKFR 47

## RESULT 11

G86403

33.3K hypothetical protein F28L5.13 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001

C:Accession: G86403

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansgen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86403

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 <STO>

A:Cross-references: GB:AE005172; NID:gl0998925; PIDN:AAG26065.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

## Query Match

Best Local Similarity 56.3%; Score 40; DB 2; Length 289;

Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
|||:|:|:|  
Db 8 CFTWEYARHVR 19

## RESULT 12

T29571

hypothetical protein F30B5.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T29571

R:Miller, N.; Bradshaw, H.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid F30B5.

A:Reference number: Z20644

A:Accession: T29571

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-361 <MIL>

A:Cross-references: EMBL:U42437; PIDN:AAA83493.1; CESP:F30B5.4

C:Genetics:

A:Gene: CESP:F30B5.4

A:Introns: 140/3; 179/2; 223/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F30B5.4

## Query Match

Best Local Similarity 56.3%; Score 40; DB 2; Length 361;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWERNMKR 10  
|||:|:|:|  
Db 8 CIQWELNRRR 17

## RESULT 13

AD2346

hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AD2346

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2346

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA76022.1; PID:gl7133459; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4323

## Query Match

Best Local Similarity 54.9%; Score 39; DB 2; Length 298;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWERNMKR 10  
|||:|:|:|  
Db 163 FHWQRYNRK 171

## RESULT 14

I37477

MHC class I histocompatibility antigen alpha chain HLA-A34 precursor - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I37477; S16771; S16767

R:Madrigal, J.A.; Bellich, M.P.; Hildebrand, W.H.; Benjamin, R.J.; Little, A.M.; Zemmour,

J. Immunol. 149, 3411-3415, 1992

A:Title: Distinctive HLA-A,B antigens of black populations formed by interallelic conver

A:Reference number: I37476; MUID:93056508; PMID:1431115

A:Accession: I37477

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-365 <RES>

A:Cross-references: EMBL:X61704; NID:g32161; PIDN:CAA43873.1; PID:g32162

A:Note: this allele is designated A\*3401 (formerly HLA-Aw34.1)  
 R;Madrigal, J.A.; Bellch, M.P.; Hildebrand, W.H.; Benjamin, R.J.; Little, A.M.; Zemmour,  
 submitted to the EMBL Data Library, August 1991  
 A:Description: Molecular definition of HLA-A,B antigens of black populations: Implication  
 A:Reference number: S16766  
 A:Accession: S16771  
 A:Molecule type: mRNA  
 A:Residues: 1-2, 'V', 4-89, 'N', 91-120, 'I', 122-128, 'S', 130-137, 'R', 139-179, 'L', 181-248, 'I',  
 A:Cross-references: EMBL:X61705  
 A:Note: this allele is designated A\*3402  
 C:Genetics:  
 A:Gene: GDB:HLA-A  
 A:Cross-references: GDB:119310; OMIM:142800  
 A:Map position: 6p21.3-6p21.3  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: heterodimer; transmembrane protein  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 54.9%; Score 39; DB 2; Length 365;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WERNMKVR 12  
 | | | | |  
 Db 84 WDRTRKVK 92

RESULT 15  
 A45687  
 outer capsid protein VP4 - human rotavirus B  
 C:Species: human rotavirus B  
 C:Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 19-Feb-1999  
 C:Accession: A45687  
 R;MacKow, E.R.; Werner-Eckert, R.; Fay, M.E.; Tao, H.; Chen, G.  
 J. Virol. 67, 2730-2738, 1993  
 A:Title: Identification and baculovirus expression of the VP4 protein of the human group  
 A:Reference number: A45687; MUID:93233240; PMID:8386274  
 A:Accession: A45687  
 A>Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-749 <MAC>  
 A:Note: sequence inconsistent with the nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:130353, NCBIP:130355)  
 C:Superfamily: rotavirus core protein

Query Match 54.9%; Score 39; DB 2; Length 749;  
 Best Local Similarity 50.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
 | | | | |  
 Db 195 CFTWDMNCANVR 206

Search completed: February 21, 2003, 07:47:56  
 Job time : 10.65 secs

GenCore version 5.1.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-84

Perfect score: 71

Sequence: 1 CFQWERNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	95.8	711	1	TRFL_HUMAN
2	51	71.8	708	1	TRFL_CAMDR
3	51	71.8	708	1	TRFL_CAPHI
4	43	60.6	695	1	TRFL_HORSE
5	42	59.2	707	1	TRFL_MOUSE
6	41	57.7	4568	1	DYHB_CHLRE
7	40	56.3	455	1	YKYL_CAEEL
8	39	54.9	365	1	1A34_HUMAN
9	39	54.9	749	1	VP4_ROTGA
10	38	53.5	146	1	RPOB_LIBAF
11	38	53.5	292	1	NLA_DROME
12	38	53.5	502	1	C91I_ARATH
13	37	52.1	62	1	RL28_THETN
14	37	52.1	152	1	YB83_METUA
15	37	52.1	334	1	CATL_MOUSE
16	37	52.1	334	1	CATL_RAT
17	37	52.1	765	1	Y008_HUMAN
18	37	52.1	783	1	YNR2_CAEEL
19	37	52.1	1179	1	ATX1_ARATH
20	37	52.1	3680	1	DMD_CANFA
21	36	50.7	329	1	CATK_RAT
22	36	50.7	333	1	CATJ_MOUSE
23	36	50.7	637	1	MUTA_PROPR
24	36	50.7	857	1	BF2_CHICK
25	36	50.7	857	1	BF2_CRIGR
26	36	50.7	857	1	BF2_HUMAN
27	36	50.7	857	1	BF2_MESAU
28	36	50.7	857	1	BF2_MOUSE
29	36	50.7	857	1	BF2_RAT
30	36	50.7	1135	1	PHYC_SORBI
31	35	49.3	160	1	Y4LN_RHISN
32	35	49.3	214	1	VIF_SIV54
33	35	49.3	388	1	5H4_HUMAN

34 35 49.3 415 1 BPS2\_HUMAN Q13515 homo sapien  
35 35 49.3 428 1 SVH\_CHIMU Q9PJ99 chlamydia m  
36 35 49.3 663 1 PDII\_HUMAN Q9ULC6 homo sapien  
37 35 49.3 1574 1 RPOC\_AQUAE Q67763 aquifex aeo  
38 35 49.3 1576 1 RPOC\_AQUAPY Q9X6V2 aquifex pyr  
39 35 49.3 2096 1 BP28\_DROME Q9VM75 drosophila  
40 34 47.9 60 1 RL28\_BACST P23374 bacillus st  
41 34 47.9 62 1 NXS3\_NAJMO P01432 naja mossam  
42 34 47.9 62 1 RL28\_LISMO Q92AJ2 listeria mo  
43 34 47.9 62 1 RL28\_STAAM Q99UP4 staphylococ  
44 34 47.9 246 1 Y495\_SYNY3 Q55185 synecocyst  
45 34 47.9 267 1 Y125\_MYCCA P53661 mycoplasma

## ALIGNMENTS

RESULT 1  
TRFL\_HUMAN  
ID TRFL\_HUMAN STANDARD; PRT; 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H123; Q96K24;  
AC Q96K25;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrroxin A;  
DE Lactoferrroxin B; Lactoferrroxin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Cho Y.Y.;  
RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Conneely O.M.;  
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Liang Q., Jimenez-Flores R., Richardson T.;  
RT "Molecular cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX Wei X., Han J., Rado T.A.;  
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences.";  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Cheng H., Chen X., Huan L.;  
RT "cDNA cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX Strausberg R.;  
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RA Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactoferrin: amino acid sequence and structural  
comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
alignment of the cyanogen bromide fragments and characterization of  
N- and C-terminal domains";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=86001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nham M., Parnell L., Dedha N., Ansari A., Mardis E., Schutz K.,  
RA Guo J., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
binding properties and crystal structure of the histidine-  
253-->methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=1009508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioind antagonist peptides derived  
from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.I.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83205; AAA59511.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; M18642; AAA66665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFFUL.  
CC PDB; 1LGF; 31-AUG-94.  
CC PDB; 1LCT; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFT; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.



Query Match 95.8%; Score 68; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0026;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
 DB 39 CFQWERNMKVR 50

RESULT 2  
 TRFL CAMDR  
 ID TRFL CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9WZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sonali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R.; Ackermann M., Farah Z., Fuhan Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin."  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION. USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AJ131674; CAB53387.1; -;  
 EMBL; AF165879; AAF82241.1; -;  
 DR HSP; O77811; 1B1X.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.

FT SIGNAL. 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 446 666 BY SIMILARITY.  
 FT DISULFID 474 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 71.8%; Score 51; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.23;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWERNMKVR 12  
 DB 38 CAQWOREMKVR 49

RESULT 3  
 TRFL CAPHI  
 ID TRFL CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Nocart M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 RT relevant locus to bovine U12 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION. USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.



FT DISULFID 463 538  
 FT DISULFID 487 681  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66 66  
 FT METAL 98 98  
 FT METAL 138 198  
 FT METAL 259 259  
 FT METAL 401 401  
 FT METAL 439 439  
 FT METAL 532 532  
 FT METAL 601 601  
 FT BINDING 127 127  
 FT BINDING 469 469  
 FT CARBOHYD 143 143  
 FT CARBOHYD 287 287  
 FT CARBOHYD 482 482  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 60.6%; Score 43; DB 1; Length 695;  
 Best Local Similarity 58.3%; Pred. No. 5.4;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWERNMKVR 12  
 Db 25 CAKFORNMKKVR 36

## RESULT 5

TRFL\_MOUSE  
 ID TRFL\_MOUSE STANDARD; PRT; 707 AA.  
 AC P08071; P70690; Q61799; Q922P2;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=87280033; PubMed=3611056;  
 RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse  
 uterine secretions.";  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Moriishi K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=32042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21890-21895(1991).  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).

CC EMBL; J03298; AAA40525.1; -  
 DR EMBL; D88510; BAA13633.1; -  
 DR EMBL; BC006904; AAH06904.1; -  
 DR EMBL; M74778; AAA39427.1; -  
 DR PIR; A28438; A28438.  
 DR HSP; P02788; ICB6.  
 DR MGD; MGI:96837; Ltf.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.  
 FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 605 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139 ANION (POTENTIAL).  
 FT BINDING 481 481 ANION (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2 MR -> IQG (IN REF. 1).  
 FT CONFLICT 25 25 R -> Q (IN REF. 2).  
 FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;

Query Match 59.2%; Score 42; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 8.1;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWERNMKVR 11  
 Db 37 CLRWQNMKKV 47

RESULT 6  
 DYHB CHLRE STANDARD; PRT; 4568 AA.  
 AC Q39565;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 39, Last annotation update)  
 DE Dynein beta chain, flagellar outer arm.  
 GN ODA4 OR ODA-4 OR SUP1.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21GT;  
 RX MEDLINE=94274778; PubMed=8006077;  
 RA Mitchell D.R., Brown K.S.;  
 RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes";  
 RL J. Cell Sci. 107:635-644(1994).  
 CC -!- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE ACTIVITY.  
 CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
 CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U02963; AAA19956.1;  
 DR InterPro; IPR004273; Dynein heavy.  
 DR Pfam; PF03028; Dynein heavy; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 KW Coiled coil.  
 FT DOMAIN 277 293 COILED COIL (POTENTIAL).  
 FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).  
 FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).  
 FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).  
 FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).  
 FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).  
 FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).  
 FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).  
 FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1919 1926 ATP (POTENTIAL).  
 FT NP\_BIND 2202 2209 ATP (POTENTIAL).  
 FT NP\_BIND 2530 2537 ATP (POTENTIAL).  
 FT NP\_BIND 2879 2886 ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;  
 Query Match 57.7%; Score 41; DB 1; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 80;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFQWERNMKVR 12  
 Db 1852 CFQWQSQLRYIQ 1863  
 RESULT 7  
 YKYL CABEL STANDARD; PRT; 455 AA.  
 ID YKYL CABEL  
 AC Q19910;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein F30B5.4 in chromosome IV.  
 GN F30B5.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Miller N., Bradshaw H.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBSJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE OXK138 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U42437; AAA83493.2;  
 DR WormPep; F30B5.4; CE28552.  
 KW Hypothetical protein.  
 SQ SEQUENCE 455 AA; 51148 MW; D8E6909408FD6C69 CRC64;  
 Query Match 56.3%; Score 40; DB 1; Length 455;  
 Best Local Similarity 60.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFQWERNWK 10  
 Db 102 CIQWELNRR 111  
 RESULT 8  
 1A34 HUMAN STANDARD; PRT; 365 AA.  
 ID 1A34 HUMAN  
 AC P30453; P30454;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE HLA class I histocompatibility antigen, AW-34(A-10) alpha chain precursor.  
 DE HLA-A OR HLA.  
 GN HLA-A OR HLA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (A\*3401/A\*3402).  
 RX MEDLINE=93056508; PubMed=1431115;  
 RA Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J., Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L., Martell R.W., du Toit E.D., Parham P.;  
 RA "Distinctive HLA-A,B antigens of black populations formed by interallelic conversion";  
 RT J. Immunol. 149:3411-3415(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (A\*3401/A\*3402).  
 RX MEDLINE=93235211; PubMed=8475492;  
 RA Madrigal J.A., Hildebrand W.H., Belich M.P., Benjamin R.J., Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L., du Toit E.D., Parham P.;  
 RA "Structural diversity in the HLA-A10 family of alleles: correlations with serology";  
 RT Tissue Antigens 41:72-80(1993).  
 RL



OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.

!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}(N).

!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA' CHAIN.

!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; U09675; AAA19557.1; -  
 InterPro; IPR001572; RNA pol B.  
 Pfam; PF00562; RNA pol B; 1.  
 PROSITE; PS01166; RNA POL BETA; PARTIAL.  
 KW Transferase; Transcription; DNA-directed RNA polymerase.  
 FT NON TER 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD6FDB943 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 146;  
 Best Local Similarity 60.0%; Pred No. 8.1;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CPQWERNRKL 10  
 |||||  
 Db 10 CVQWSRGARK 19  
 |||||

RESULT 11  
 NLA DROME  
 ID NLA DROME STANDARD; PRT; 292 AA.  
 AC Q9XZL8; Q9V391;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 OS NLA OR CG6072.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP McCormick A.V., Goldberg M.L.;  
 RT "Gene required for elongation of meiosis I spindle in Drosophila females."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Berkley;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Benson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glöckle A., Gong P., Gorrell J.H.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svavarskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 CC !- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 CC !- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF147700; AAD33987.1; -  
 DR EMBL; AE003712; AAF55285.1; -  
 DR FlyBase; FBgn026629; nla.  
 SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 292;  
 Best Local Similarity 54.5%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FQWERNRKL 12  
 |||||  
 Db 150 FQWERSFRLLR 160  
 |||||

RESULT 12  
 C911 ARATH  
 ID C911 ARATH STANDARD; PRT; 502 AA.  
 AC Q9FG65; O65789;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cytochrome P450 91A1 (EC 1.14.-.-).  
 CN CYP91A1 OR AT5G36220 OR T30G6.3.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=cv. Columbia;  
 RC MEDLINE=21016721; PubMed=11130714;  
 RX Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asanizu E.,  
 RA Miyajima N., Saezamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,

RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,  
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,  
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowatt C.,  
 RA Wagner-McPherson C., Wollam A., Yakum M., Bell M., Dedhia N.,  
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,  
 RA Kirschoff K., Toth K., King L., Bahret A., Miller B., Marra M.,  
 RA Martienssen R., McComble W.R., Wilson R.K., Murphy G., Bancroft I.,  
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.-A., McCallagh B., Robben J., Grymouprez B., Zimmermann W.,  
 RA Ramsberger U., Wedler H., Balke K., Wedler E., Peters S.,  
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,  
 RA Weitzsaecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
 RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,  
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,  
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,  
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:823-826(2000).  
 RN [2]  
 RP SEQUENCE OF 3-502 FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98281573; PubMed=9620263;  
 RA Mizutani M., Ward E., Ohta D.;  
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of  
 RT cDNAs, differential expression, and RFLP mapping of multiple  
 RT cytochromes P450";  
 RL Plant Mol. Biol. 37:39-52(1998).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB026661; BAB09361.1; -;  
 DR EMBL; D78606; BAA28538.1; -;  
 DR HSSP; P14779; IJPZ.  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.  
 FT TRANSMEM 6 26 POTENTIAL.  
 FT BINDING 440 440 HEME (BY SIMILARITY).  
 SQ SEQUENCE 502 AA; 56723 MW; 1A348A8DD76A41B0 CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 502;  
 Best Local Similarity 83.3%; Pred. No. 28;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CPQWER 6  
 DB 460 CFEWER 465  
 ||:||||  
 RESULT 13  
 RL28\_THETN  
 ID RL28\_THETN STANDARD; PRT; 62 AA.  
 AC Q8R9U1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L28.  
 GN RPBH OR TTE1495.  
 OS Thermoanaerobacter tengcongensis.  
 CC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 CC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 NCBI\_TaxID=119072;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome";  
 RL Genome Res. 12:689-700(2002).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE013107; AAM24713.1; -;  
 DR Ribosomal protein; Complete proteome.  
 KW SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 Query Match 52.1%; Score 37; DB 1; Length 62;  
 Best Local Similarity 60.0%; Pred. No. 5.1;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QWERNRKVR 12  
 DB 27 RWKPNIRKVR 36  
 ||:||||  
 RESULT 14  
 YE83\_METJA  
 ID YE83\_METJA STANDARD; PRT; 152 AA.  
 AC Q58878;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MJ1483.  
 GN MJ1483.  
 OS Methanococcus jannaschii.  
 CC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 CC Methanocaldococcaceae; Methanocaldococcus.  
 NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RA MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii";  
 RL Science 273:1058-1073(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U67589; AAB99498.1; -;  
 DR TIGR; MJ1483; -;  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 7 27 POTENTIAL.

```

SQ SEQUENCE 152 AA; 17659 MW; 90227132448B8802 CRC64;
Query Match 52.1%; Score 37; DB 1; Length 152;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 FQWERNMKRV 11
DB 68 FDWERNINNI 77
-----
RESULT 15
CATL_MOUSE
ID CATL_MOUSE STANDARD; PRT; 334 AA.
AC P06797;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin L precursor (EC 3.4.22.15) (Major excreted protein) (MEP).
GN CTSL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88076849; PubMed=3689328;
RA Troen B.R., Gal S., Gottesman M.M.;
RT "Sequence and expression of the cDNA for MEP (major excreted
RL protein), a transformation-regulated secreted cathepsin.";
RN Biochem. J. 246:731-735(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88213715; PubMed=2835398;
RA Joseph L.J., Chang L.C., Stamenkovich D., Sukhatme V.P.;
RT "Complete nucleotide and deduced amino acid sequences of human and
RL murine preprocathepsin L. An abundant transcript induced by
RN transformation of fibroblasts.";
RX J. Clin. Invest. 81:1621-1629(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87033683; PubMed=3533924;
RA Portnoy D.A., Erickson A.H., Kochan J., Ravetch J.V., Unkless J.C.;
RT "Cloning and characterization of a mouse cysteine proteinase.";
RN J. Biol. Chem. 261:14697-14703(1986).
RN [4]
RP SEQUENCE FROM N.A., AND CARBOHYDRATE-LINKAGE SITE.
RC TISSUE=Liver;
RX MEDLINE=91112761; PubMed=2275556;
RA Stearns N.A., Dong J., Pan J.X., Brenner D.A., Sahagian G.G.;
RT "Comparison of cathepsin L synthesized by normal and transformed
RL cells at the gene, message, protein, and oligosaccharide levels.";
RN Arch. Biochem. Biophys. 283:447-457(1990).
RN [5]
RP SEQUENCE OF 89-300 FROM N.A.
RC STRAIN=BNL;
RX MEDLINE=86271744; PubMed=3755373;
RA Denhardt D.T., Hamilton R.T., Parfett C.L.J., Edwards D.R.,
RA Pierre R.S., Waterhouse P., Nilsson-Hamilton M.;
RT "Close relationship of the major excreted protein of transformed
RL murine fibroblasts to thiol-dependent cathepsins.";
RN Cancer Res. 46:4590-4593(1986).
CC -!- FUNCTION: IMPORTANT FOR THE OVERALL DEGRADATION OF PROTEINS IN
CC LYOSOMES.
CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain. As
CC compared to cathepsin B, cathepsin L exhibits higher activity
CC towards protein substrates, but has little activity on Z-Arg-Arg-
CC NHec, and no peptidyl-dipeptidase activity.
CC -!- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE
CC BONDS.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06086; CAA29470.1; -
DR EMBL; J02583; AAA37445.1; -
DR EMBL; M20495; AAA39984.1; -
DR EMBL; X04392; CAA27980.1; -
DR PIR; S01177; KMSL.
DR PIR; S13890; S13890.
DR HSP; P07711; ICUL.
DR MEROPS; C01.032; -.
DR MGD; MGI:89564; Ctsl.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR00169; SHprot_acSite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR PRODOM; PD000158; Peptidase_C1; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
FT SIGNAL 1 17
FT PROPEP 18 113
FT CHAIN 114 288
FT CHAIN 291 334
FT ACT_SITE 138 138
FT ACT_SITE 276 276
FT ACT_SITE 300 300
FT DISULFID 135 178
FT DISULFID 169 211
FT CARBOHYD 221 221
FT CONFLICT 58 58
FT CONFLICT 177 177
SQ SEQUENCE 334 AA; 37547 MW; FE6747043307AD98 CRC64;
Query Match 52.1%; Score 37; DB 1; Length 334;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 WERNMKRV 12
DB 52 WEKNRMQ 60
-----
Search completed: February 21, 2003, 07:27:59
Job time : 5.6 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107b-84

Perfect score: 71

Sequence: 1 CFQWERNMRKV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	88.7	711	4	Q8TCD2
2	59	83.1	38	4	Q9UCY5
3	48	67.6	33	6	Q9TR80
4	44	62.0	511	16	Q8Z462
5	43	60.6	279	16	Q8XSE2
6	42	59.2	105	10	Q9XFD5
7	42	59.2	306	4	Q8TAX2
8	42	59.2	466	4	Q9NUS2
9	41	57.7	205	8	Q98R82
10	41	57.7	275	5	Q93780
11	41	57.7	372	10	Q81653
12	41	57.7	1432	10	Q23230
13	41	57.7	2186	5	Q9N906
14	40	56.3	91	15	Q77856
15	40	56.3	115	15	Q9Q910
16	40	56.3	289	10	Q9C6N2

```

17 39 54.9 148 10 Q9XHP1
18 39 54.9 298 16 Q8YF77
19 39 54.9 332 11 Q9DAZ8
20 39 54.9 332 11 Q91ZD5
21 39 54.9 333 11 Q91Z75
22 39 54.9 363 7 Q31215
23 39 54.9 368 7 Q19441
24 39 54.9 820 16 Q9KRF0
25 38.5 54.2 395 12 Q92283
26 38 53.5 91 15 Q77855
27 38 53.5 185 10 Q9SM68
28 38 53.5 273 2 Q31090
29 38 53.5 274 4 Q96M21
30 38 53.5 341 11 Q8R2A4
31 38 53.5 393 16 Q9ZKP4
32 38 53.5 433 5 Q44473
33 38 53.5 481 10 Q9C714
34 38 53.5 489 16 P96223
35 38 53.5 492 10 Q23155
36 38 53.5 499 10 Q9SZT6
37 38 53.5 500 10 Q23154
38 38 53.5 509 10 Q9LHA1
39 38 53.5 570 10 Q8S487
40 38 53.5 584 17 Q9HPA3
41 38 53.5 887 10 Q9SX85
42 38 53.5 932 5 Q19153
43 37 52.1 62 16 Q8R9U1
44 37 52.1 109 15 Q9YQC1
45 37 52.1 109 15 Q9YQC0

```

#### ALIGNMENTS

```

RESULT 1
Q8TCD2 PRELIMINARY; PRT; 711 AA.
ID Q8TCD2;
AC Q8TCD2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Lactotransferrin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RL Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022347; AAH22347.1; -.
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

```

Query Match 88.7%; Score 63; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 CFQWERNMRKV 11
Db 39 CFQWERNMRKV 49

```

```

RESULT 2
Q9UCY5 PRELIMINARY; PRT; 38 AA.
ID Q9UCY5
AC Q9UCY5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE Lactoferrin homolog (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96081613; PubMed=8551695;
RA Sato I.;
RT "Characterization of the 84-kDa protein with ABH activity in human
  seminal plasma.";
RL Jpn. J. Legal Med. 49:281-293 (1995).
DR HSP; P02788; 1BXA.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDDDB CRC64;

Query Match      83.1%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.00085;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FQWERNMKVR 12
Db 21 FQWERNMKVR 31
|||:|||||
|||:|||||

RESULT 3
Q9TR80
ID Q9TR80 PRELIMINARY; PRT; 33 AA.
AC Q9TR80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Lactoferrin (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=95127729; PubMed=7827104;
RA Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;
RL Biochim. Biophys. Acta 1243:25-32 (1995).
DR HSP; O77698; 1CE2.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match      67.6%; Score 48; DB 6; Length 33;
Best Local Similarity 54.5%; Pred. No. 0.076;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWERNMKV 11
Db 19 CYQWQKMKRL 29
|||:|||||
|||:|||||

RESULT 4
Q8Z462
ID Q8Z462 PRELIMINARY; PRT; 511 AA.
AC Q8Z462;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein SRY3070.
GN STY3070.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
SQ MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

```

```

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
DR EMBL; AL627276; CAD06049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DDDD124E10D178B CRC64;

Query Match      62.0%; Score 44; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWERNMKVR 12
Db 350 CFQWERNMKVR 361
|||:|||||
|||:|||||

RESULT 5
Q8XSE2
ID Q8XSE2 PRELIMINARY; PRT; 279 AA.
AC Q8XSE2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative ICC protein homolog.
GN ICC OR RSF0334 OR RS00414.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler C., Choisme N., Claudel-Renard C., Cumnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AL646079; CAD17685.1; -.
DR InterPro; IPR004843; M-ppestrase.
DR Pfam; PF00149; Metallophos; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 279 AA; 31541 MW; ABB38818004B2EDA CRC64;

Query Match      60.6%; Score 43; DB 16; Length 279;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWERNMKVR 12
Db 244 CFQWERNMKVR 255
|||:|||||
|||:|||||

RESULT 6
Q9XFD5
ID Q9XFD5 PRELIMINARY; PRT; 105 AA.
AC Q9XFD5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cytochrome P450 (Fragment).

```

OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoidae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;

[1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=PANICLE;  
 RA Liu J., Yang J.;  
 RT "Suppression subtractive hybridization (SSH) identified candidate  
 RT genes that are differentially expressed at rice young panicle."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL: AF140486; AAD29699.1; -  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 FT NON\_TER 1  
 SQ SEQUENCE 105 AA; 11912 MW; B0EFCD487E19F9 CRC64;

Query Match 59.2%; Score 42; DB 10; Length 105;  
 Best Local Similarity 70.0%; Pred. No. 3.2;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CPOWERMRK 10  
 ||||| :  
 DB 61 CPOWERLGKK 70

RESULT 7  
 Q8TAX2 PRELIMINARY; PRT; 306 AA.  
 ID Q8TAX2;  
 AC Q8TAX2;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to hypothetical protein FLJ11175.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANCREAS;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC025708; AAH25708.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 306 AA; 34551 MW; 43D9AD722566B343 CRC64;

Query Match 59.2%; Score 42; DB 4; Length 306;  
 Best Local Similarity 66.7%; Pred. No. 10;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CPOWERMR 9  
 ||||| :  
 DB 269 CFQWESTLR 277

RESULT 8  
 Q9NUS2 PRELIMINARY; PRT; 466 AA.  
 ID Q9NUS2;  
 AC Q9NUS2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE CDNA FLJ11175 fis, clone P1ACB100735, weakly similar to phorbol  
 DE ester/diacylglycerol-binding protein UNC-13.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,  
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Takiguchi S.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Yamamoto J., Wakamatsu A., Nakamori Y., Saito K.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK002037; BAA92048.1; -  
 DR HSSP: P21707; IBYN.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR000504; RNA\_rec\_mot.  
 DR Pfam: PF00168; C2; 1.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR SMART: SM00239; C2; 1.  
 DR PROSITE: PS00499; C2\_DOMAIN 1; UNKNOWN\_1.  
 DR PROSITE: PS00004; C2\_DOMAIN 2; 1.  
 DR PROSITE: PS00030; RRM\_RNP 1; UNKNOWN\_1.  
 SQ SEQUENCE 466 AA; 53192 MW; E4113A5062F58D6E CRC64;

Query Match 59.2%; Score 42; DB 4; Length 466;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CPOWERMR 9  
 ||||| :  
 DB 269 CFQWESTLR 277

RESULT 9  
 Q98RR2 PRELIMINARY; PRT; 205 AA.  
 ID Q98RR2;  
 AC Q98RR2;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE 26S proteasome SU B5.  
 GN PRB5.  
 OS Guillardia theta (Cryptomonas phi).  
 OG Nucleomorph.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21223349; PubMed=11323671;  
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,  
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;  
 RT "The highly reduced genome of an enslaved algal nucleus."  
 RL Nature 410:1091-1096 (2001).  
 DR EMBL: AF165818; AAK39885.1; -  
 DR InterPro: IPR000243; Proteasome\_B.  
 DR InterPro: IPR001353; Proteasome.  
 DR Pfam: PF00227; Proteasome; 1.  
 DR PRINTS: PR00141; PROTEASOME.  
 KW Proteasome.  
 SQ SEQUENCE 205 AA; 22691 MW; D30F5289CBC85049 CRC64;

Query Match 57.7%; Score 41; DB 8; Length 205;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CPOWERMR 8  
 ||||| :  
 DB 63 CFFWRNL 70

RESULT 10  
 Q93780

ID Q93780 PRELIMINARY; PRT; 275 AA.  
 AC Q93780;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE F53H4.4 protein.  
 GN F53H4.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dobson R.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81089; CAB03137.1;  
 SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;  
 Query Match 57.7%; Score 41; DB 5; Length 275;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOWERNMKVR 12  
 : : : : :  
 Db 262 FOWKLSMRKTR 272

RESULT 11  
 ID 081653 PRELIMINARY; PRT; 372 AA.  
 AC 081653;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Senescence-associated protein 3 (Fragment).  
 GN SA3.  
 OS Hemerocallis hybrid cultivar.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
 OC Hemerocallidaceae; Hemerocallis.  
 OX NCBI\_TaxID=80862;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. STELLA D'ORO; TISSUE=SENESCING PETALS;  
 RX MEDLINE=99339248; PubMed=10412903;  
 RA Panavas T., Pikula A., Reid P.D., Rubinstein B., Walker E.L.;  
 RT "Identification of senescence-associated genes from daylily petals."  
 RT Plant Mol. Biol. 40:237-248(1999).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF082028; AAC34853.1;  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 FT NON TER 1  
 SQ SEQUENCE 372 AA; 41909 MW; 55FB3EAF9695955E CRC64;  
 Query Match 57.7%; Score 41; DB 10; Length 372;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWERN 7  
 : : : : :  
 Db 328 CFOWERN 334

RESULT 12  
 ID 023230 PRELIMINARY; PRT; 1432 AA.  
 AC 023230;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Trichoyalin like protein.  
 GN C7A10.840 OR AT4G36520.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Schueller C.,  
 RA Chawatzis N.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z99708; CAB16841.1;  
 DR EMBL; AL161589; CAB80318.1;  
 SQ SEQUENCE 1432 AA; 163973 MW; 96CE6FEBCB9BF0538 CRC64;  
 Query Match 57.7%; Score 41; DB 10; Length 1432;  
 Best Local Similarity 54.5%; Pred. No. 79;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOWERNMKVR 12  
 : : : : :  
 Db 583 YDWEQNARKLR 593

RESULT 13  
 ID Q9N906 PRELIMINARY; PRT; 2186 AA.  
 AC Q9N906;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Conserved hypothetical protein CHRI.67.  
 GN CHRI.67.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TREU927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerrard C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL359782; CAB95377.1;  
 DR InterPro; IPR002035; VWFA.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 2186 AA; 245737 MW; 78BB75505012005A CRC64;  
 Query Match 57.7%; Score 41; DB 5; Length 2186;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWERN 7  
 : : : : :  
 Db 588 CYEWEERN 594

RESULT 14

Search completed: February 21, 2003, 07:44:38  
Job time : 21.8 secs

```
Q77856
ID Q77856 PRELIMINARY; PRT; 91 AA.
AC Q77856;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
EN Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95191002; PubMed=7884875;
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
RA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
RT "Similarity in env and gag genes between genomic RNAs of human
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is
RT unrelated to time of HIV-1 RNA positivity in the child.";
RL J. Virol. 69:2285-2296(1995).
DR EMBL; 247868; CAAB7882.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 91
SQ SEQUENCE 91 AA; 10519 MW; EB20C63A22DA1288 CRC64;

Query Match 56.3%; Score 40; DB 15; Length 91;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWERNMKVR 12
DB 67 QWDRTLQKVR 76
|||:|:|

RESULT 15
Q9Q9L0
ID Q9Q9L0 PRELIMINARY; PRT; 115 AA.
AC Q9Q9L0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
EN Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VTRJ07;
RA Tanuri A., Machado E., Caride E., Costa L.J., Telles J.G.;
RA "Primary infections with HIV-1 of women and their offspring in Rio de
RT Janeiro, Brazil: Finding of recombinant of HIV-1 subtypes B and F.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF153457; AAF16896.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 12832 MW; 86E509D6F6978127 CRC64;

Query Match 56.3%; Score 40; DB 15; Length 115;
Best Local Similarity 60.0%; Pred. No. 8.3;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWERNMKVR 12
DB 61 QWERTLQKVR 70
|||:|:|
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 Seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107B-85  
Perfect score: 71  
Sequence: 1 CFQWQENRKRVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	12	21	AA198085 Human lactoferrin
2	67	94.4	12	21	AA198081 Human lactoferrin
3	66	93.0	12	21	AA198038 Human lactoferrin
4	66	93.0	12	21	AA198046 Human lactoferrin
5	66	93.0	12	21	AA198047 Human lactoferrin
6	66	93.0	13	21	AA198037 Human lactoferrin
7	66	93.0	13	21	AA198048 Human lactoferrin
8	66	93.0	13	21	AA198049 Human lactoferrin
9	66	93.0	14	21	AA198036 Human lactoferrin
10	66	93.0	14	21	AA198050 Human lactoferrin

11	66	93.0	14	21	AA198051 Human lactoferrin
12	66	93.0	15	17	AA198554 Peptide for anti-u
13	66	93.0	15	21	AA198035 Human lactoferrin
14	66	93.0	15	21	AA198062 Human lactoferrin
15	66	93.0	15	21	AA198063 Human lactoferrin
16	66	93.0	16	21	AA198031 Human lactoferrin
17	66	93.0	16	21	AA198064 Human lactoferrin
18	66	93.0	16	21	AA198065 Human lactoferrin
19	66	93.0	17	21	AA198034 Human lactoferrin
20	66	93.0	17	21	AA198066 Human lactoferrin
21	66	93.0	17	21	AA198067 Human lactoferrin
22	66	93.0	18	15	AA198552 Advanced glycosyla
23	66	93.0	18	17	AA198397 Human lactoferrin
24	66	93.0	18	21	AA198033 Amino acid sequenc
25	66	93.0	19	21	AA198067 Human lactoferrin
26	66	93.0	19	21	AA198032 Anti microbial pep
27	66	93.0	20	13	AA198180 Lactoferrin-relate
28	66	93.0	20	14	AA198481 Lactoferrin derive
29	66	93.0	20	15	AA198530 Lactoferrin derive
30	66	93.0	20	15	AA198531 Lactoferrin derive
31	66	93.0	20	15	AA198546 Bovine lactoferrin
32	66	93.0	20	16	AA198498 Bovine lactoferrin
33	66	93.0	20	16	AA198499 Anti-parasitic lac
34	66	93.0	20	16	AA198523 Anti-parasitic lac
35	66	93.0	20	16	AA198524 Peptide for anti-u
36	66	93.0	20	17	AA198553 Lactoferrin-derive
37	66	93.0	20	17	AA198552 Lactoferrin-derive
38	66	93.0	20	17	AA198552 Lactoferrin-derive
39	66	93.0	20	17	AA198552 Lactoferrin-derive
40	66	93.0	20	17	AA198552 Lactoferrin-derive
41	66	93.0	20	17	AA198552 Lactoferrin-derive
42	66	93.0	20	17	AA198552 Lactoferrin-derive
43	66	93.0	20	18	AA198552 Lactoferrin-derive
44	66	93.0	20	18	AA198552 Lactoferrin-derive
45	66	93.0	20	19	AA198552 Lactoferrin-derive

ALIGNMENTS

RESULT 1  
AA198085  
ID AA198085 standard; Peptide; 12 AA.  
XX AA198085;  
AC  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:85.  
DE  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
OS  
XX WO200001730-A1.  
FN  
XX  
XX 13-JAN-2000.  
PD  
XX  
XX 06-JUL-1999; 99WO-SE01230.  
PF  
XX  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
XX (ASCI-) A+ SCI INVEST AB.  
PA  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI  
XX WPI; 2000-147388/13.  
DR

XX New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
Query Match 100.0%; Score 71; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
Db 1 CFQWQENMRKVR 12

RESULT 2  
AAY78081  
ID AAY78081 standard; Peptide; 12 AA.  
XX  
AC AAY78081;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:81.  
XX  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;  
Query Match 94.4%; Score 67; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
Db 1 CFQWQENMRKVR 12

RESULT 3  
AAY78038  
ID AAY78038 standard; Peptide; 12 AA.  
XX  
AC AAY78038;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:38.  
XX  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 12; Page 70; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00016;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
 DB 1 CFQWQENMRKVR 12  
 ||||| |||||

RESULT 4  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 AC AAY78046;  
 DT 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:46.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SF01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00016;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
 DB 1 CFQWQENMRKVR 12  
 ||||| |||||

RESULT 5  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.  
 XX AC AAY78047;  
 XX DT 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:47.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SF01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00016;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



OY 1 CFQWQENMKVR 12  
Db 1 CFQWQENMKVR 12

## RESULT 6

AAV78037  
ID AAV78037 standard; Peptide; 13 AA.

XX AC AAV78037;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:37.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.

XX OS Homo sapiens.  
XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
XX CC lactoferrin. The peptides are taken up in the intestine through  
XX CC binding to specific lactoferrin receptors and are then transported  
XX CC through the circulation. A medicinal product of the peptide or fragment  
XX CC can be used for treating and/or prevention of infections (such as  
XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
XX CC membrane), inflammations and/or tumors. The peptides can also be used  
XX CC in food stuffs such as infant formula food. The peptides are also  
XX CC fungicidal and bactericidal and may also be used as preservatives.  
XX CC Even though native human lactoferrin have been shown to have desired  
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
XX CC cannot be used clinically on a broad basis because of high production  
XX CC costs. Therefore, provision of peptides based on lactoferrin would  
XX CC enable them to be used for the same purposes as lactoferrin at lower  
XX CC cost.

XX SQ Sequence 13 AA;

Query Match 93.0%; Score 66; DB 21; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.00018;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWQENMKVR 12

Db 2 CFQWQENMKVR 13

## RESULT 7

AAV78048

ID AAV78048 standard; Peptide; 13 AA.

XX AC AAV78048;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:48.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 74; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human  
XX CC lactoferrin. The peptides are taken up in the intestine through  
XX CC binding to specific lactoferrin receptors and are then transported  
XX CC through the circulation. A medicinal product of the peptide or fragment  
XX CC can be used for treating and/or prevention of infections (such as  
XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
XX CC membrane), inflammations and/or tumors. The peptides can also be used  
XX CC in food stuffs such as infant formula food. The peptides are also  
XX CC fungicidal and bactericidal and may also be used as preservatives.  
XX CC Even though native human lactoferrin have been shown to have desired  
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
XX CC cannot be used clinically on a broad basis because of high production  
XX CC costs. Therefore, provision of peptides based on lactoferrin would  
XX CC enable them to be used for the same purposes as lactoferrin at lower  
XX CC cost.

XX SQ Sequence 13 AA;

Query Match 93.0%; Score 66; DB 21; Length 13;  
Best Local Similarity 91.7%; Pred. No. 0.00018;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWQENMKVR 12

Db 2 CFQWQENMKVR 13

## RESULT 8

AAV78049

ID AAV78049 standard; Peptide; 13 AA.

XX AC AAV78049;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:49.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX XX (ASCI-) A+ SCI INVEST AB.  
 XX PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX PI WPI; 2000-147388/13.  
 XX DR New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 18; Page 74; 102pp; English.  
 XX XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 13 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.00018;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQENMKVR 12  
 DB 2 CFQWQENMKVR 13  
 RESULT 9  
 AAY78036  
 ID AAY78036 standard; Peptide; 14 AA.  
 XX AC AAY78036;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:36.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.

OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX XX (ASCI-) A+ SCI INVEST AB.  
 XX PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX PI WPI; 2000-147388/13.  
 XX DR New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 69; 102pp; English.  
 XX XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00019;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQENMKVR 12  
 DB 3 CFQWQENMKVR 14  
 RESULT 10  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AC AAY78050;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:50.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 75; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00019;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQENMKVR 12  
 Db 3 CFQWQENMKVR 14  
 RESULT 11  
 ID AAY78051 standard; Peptide; 14 AA.  
 AC AAY78051;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:51.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 75; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00019;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQENMKVR 12  
 Db 3 CFQWQENMKVR 14  
 RESULT 12  
 ID AAR98554 standard; Peptide; 15 AA.  
 AC AAR98554;  
 XX 12-NOV-1996 (first entry)  
 XX Peptide for anti-ulcer agent.  
 XX anti-ulcer agent; low toxicity; stable; heat-resistant.  
 XX Synthetic.  
 XX JP08143468-A.  
 XX 04-JUN-1996.  
 XX 17-NOV-1994; 94JP-0283869.  
 XX 17-NOV-1994; 94JP-0283869.  
 XX (MORG) MORINAGA MILK IND CO LTD.  
 XX WPI; 1996-318857/32.  
 XX Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble  
 XX Claim 1; Page 11; 11pp; Japanese.  
 XX AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.  
 XX SQ Sequence 15 AA;

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 75; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 14 AA;  
 Query Match 93.0%; Score 66; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00019;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQENMKVR 12  
 Db 3 CFQWQENMKVR 14  
 RESULT 12  
 ID AAR98554 standard; Peptide; 15 AA.  
 AC AAR98554;  
 XX 12-NOV-1996 (first entry)  
 XX Peptide for anti-ulcer agent.  
 XX anti-ulcer agent; low toxicity; stable; heat-resistant.  
 XX Synthetic.  
 XX JP08143468-A.  
 XX 04-JUN-1996.  
 XX 17-NOV-1994; 94JP-0283869.  
 XX 17-NOV-1994; 94JP-0283869.  
 XX (MORG) MORINAGA MILK IND CO LTD.  
 XX WPI; 1996-318857/32.  
 XX Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble  
 XX Claim 1; Page 11; 11pp; Japanese.  
 XX AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.  
 XX SQ Sequence 15 AA;

Query Match 93.0%; Score 66; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.0002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
 ||||| |||||  
 DB 2 CFQWQENMRKVR 13

## RESULT 13

AAV78035  
 ID AAV78035 standard; Peptide; 15 AA.

XX  
 AC AAV78035;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:35.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.

XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.

XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 69; 102pp; English.

XX  
 CC AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides are also  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 15 AA;

Query Match 93.0%; Score 66; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.0002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
 ||||| |||||  
 DB 4 CFQWQENMRKVR 15

## RESULT 14

AAV78062  
 ID AAV78062 standard; Peptide; 15 AA.

XX  
 AC AAV78062;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:62.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX  
 OS Homo sapiens.  
 OS Synthetic.

XX  
 PN WO200001730-A1.

XX  
 PD 13-JAN-2000.

XX  
 PF 06-JUL-1999; 99WO-SE01230.

XX  
 PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX  
 DR WPI; 2000-147388/13.

XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX  
 PS Claim 15; Page 81; 102pp; English.

XX  
 CC AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides are also  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 15 AA;

Query Match 93.0%; Score 66; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.0002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
 ||||| |||||  
 DB 4 CFQWQENMRKVR 15

## RESULT 15

AAV78063  
 ID AAV78063 standard; Peptide; 15 AA.

XX  
 AC AAV78063;

```

XX 25-APR-2000 (first entry)
XX
XX DE Human lactoferrin derived peptide SEQ ID NO:63.
XX
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX urinary tract infection; colitis; Candida infection; fungicidal;
XX bactericidal; preservative.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX PN WO200001730-A1.
XX
XX PD 13-JAN-2000.
XX
XX PF 06-JUL-1999; 99WO-SE01230.
XX
XX PR 06-JUL-1999; 98SE-0002441.
XX PR 17-JUL-1999; 98SE-0002562.
XX PR 29-DEC-1999; 98SE-0004614.
XX
XX PA (ASCI-) A+ SCI INVEST AB.
XX
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX WPI; 2000-147388/13.
XX
XX PT New peptides used for treatment and prevention of infections,
XX inflammations and tumors and for use in infant formula food
XX
XX PS Claim 18; Page 81; 102pp; English.
XX
XX AAAY78001 to AAAY78100 represent peptides having sequences based on human
XX lactoferrin. The peptides are taken up in the intestine through
XX binding to specific lactoferrin receptors and are then transported
XX through the circulation. A medicinal product of the peptide or fragment
XX can be used for treating and/or prevention of infections (such as
XX urinary tract infections, colitis, and Candida infection on a mucosal
XX membrane), inflammations and/or tumors. The peptides can also be used
XX in food stuffs such as infant formula food. The peptides are also
XX fungicidal and bactericidal and may also be used as preservatives.
XX Even though native human lactoferrin have been shown to have desired
XX anti-inflammatory anti-infectious and anti-tumoural properties they
XX cannot be used clinically on a broad basis because of high production
XX costs. Therefore, provision of peptides based on lactoferrin would
XX enable them to be used for the same purposes as lactoferrin at lower
XX cost.
XX
XX SQ Sequence 15 AA;
    Query Match 93.0%; Score 66; DB 21; Length 15;
    Best Local Similarity 91.7%; Pred. No. 0.0002;
    Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
    QY 1 CFQWQENMRKVR 12
    DB 4 CFQWQENMRKVR 15

```

Search completed: February 21, 2003, 07:37:15  
 Job time : 28.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.\*

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-85

Perfect score: 71

Sequence: 1 CFQWQENMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	93.0	18	1	US-08-204-487-3
2	66	93.0	18	2	US-08-485-948-8
3	66	93.0	18	2	US-08-628-380-8
4	66	93.0	18	2	US-08-475-055-8
5	66	93.0	20	1	US-07-755-161A-3
6	66	93.0	20	1	US-07-891-174-3
7	66	93.0	20	1	US-08-204-487-1
8	66	93.0	20	1	US-08-256-771-24
9	66	93.0	20	1	US-08-256-771-25
10	66	93.0	20	1	US-08-381-984-24
11	66	93.0	20	1	US-08-381-984-25
12	66	93.0	22	4	US-09-508-734-4
13	66	93.0	24	4	US-09-508-734-6
14	66	93.0	25	1	US-07-755-161A-10
15	66	93.0	25	1	US-07-891-174-10
16	66	93.0	25	1	US-08-204-487-7
17	66	93.0	29	4	US-09-508-734-8
18	66	93.0	36	1	US-07-755-161A-8
19	66	93.0	36	1	US-07-891-174-8
20	66	93.0	36	1	US-08-256-771-30
21	66	93.0	36	1	US-08-381-984-29
22	66	93.0	47	2	US-08-464-182A-6
23	66	93.0	47	2	US-08-406-271-6
24	66	93.0	50	2	US-08-693-274A-7
25	66	93.0	52	4	US-09-017-043A-3
26	66	93.0	53	2	US-08-464-182A-5
27	66	93.0	53	2	US-08-406-271-5

28	66	93.0	54	2	US-08-464-182A-2
29	66	93.0	54	2	US-08-406-271-2
30	66	93.0	694	3	US-08-724-586-2
31	66	93.0	694	4	US-09-421-632-2
32	66	93.0	694	4	US-09-932-190-2
33	66	93.0	705	2	US-08-655-640-2
34	66	93.0	708	2	US-08-655-640-4
35	66	93.0	711	1	US-08-154-019-4
36	66	93.0	711	1	US-08-461-333-4
37	66	93.0	711	3	US-08-464-167-4
38	66	93.0	711	3	US-09-158-313-4
39	66	93.0	711	4	US-08-476-798-4
40	63	88.7	711	1	US-08-145-681-2
41	63	88.7	711	1	US-08-250-308-2
42	63	88.7	711	1	US-08-453-703-2
43	63	88.7	711	2	US-08-456-106-2
44	63	88.7	711	3	US-08-456-108-2
45	63	88.7	711	4	US-09-265-577-2

## ALIGNMENTS

### RESULT 1

US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERU  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 93.0%; Score 66; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0;

QY 1 CFQWQENMKVR 12  
||| ||| ||| ||| |||  
Db 1 CFQWQENMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-08-485-948-8

Query Match 93.0%; Score 66; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0;

QY 1 CFQWQENMKVR 12  
||| ||| ||| ||| |||  
Db 1 CFQWQENMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-08-628-380-8

Query Match 93.0%; Score 66; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12  
||| ||| ||| ||| |||  
Db 1 CFQWQENMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-Cl, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; US-08-475-055-8

Query Match 93.0%; Score 66; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENNRKVR 12  
Db 1 CFQWQENNRKVR 12

## RESULT 5

US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
;; US-07-755-161A-3

Query Match

93.0%; Score 66; DB 1; Length 20;





```

; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
; OTHER INFORMATION: DERIVED FROM HUMAN LACTOFERRIN"
US-08-204-487-1

Query Match 93.0%; Score 66; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12
Db 2 CFQWQENMRKVR 13

RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5658591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; PRODUCT'S THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"
US-08-204-487-1

Query Match 93.0%; Score 66; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12
Db 2 CFQWQENMRKVR 13

RESULT 9
US-08-256-771-25
; Sequence 25, Application US/08256771
; Patent No. 5658591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; PRODUCT'S THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are protected to
; OTHER INFORMATION: prevent disulfide bond"
US-08-256-771-25

Query Match 93.0%; Score 66; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12
Db 2 CFQWQENMRKVR 13

RESULT 10
US-08-381-984-24
```

Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
and 19 are bonded by disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
peptides including the specified peptide as a fragment thereof"  
US-08-381-984-24  
Query Match 93.0%; Score 66; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 6e-05; 1; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQNMKVR 12  
DB 2 CFQWQNMKVR 13  
RESULT 11  
US-08-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
and 19 are chemically modified to prevent disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
peptides including the specified peptide as a fragment thereof"  
US-08-381-984-25  
Query Match 93.0%; Score 66; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 6e-05; 1; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQNMKVR 12  
DB 2 CFQWQNMKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
useful microorganism thereof  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatent in 1.71  
SEQ ID NO 4

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
peptides including the specified peptide as a fragment thereof"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
and 19 are chemically modified to prevent disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
peptides including the specified peptide as a fragment thereof"  
US-08-381-984-25  
Query Match 93.0%; Score 66; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 6e-05; 1; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQNMKVR 12  
DB 2 CFQWQNMKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
useful microorganism thereof  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatent in 1.71  
SEQ ID NO 4

LENGTH: 22  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-4

Query Match 93.0%; Score 66; DB 4; Length 22;  
Best Local Similarity 91.7%; Pred. No. 6.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
||| ||| ||| ||| |||  
Db 2 CFQWQENMRKVR 13

## RESULT 13

US-09-508-734-6  
Sequence 6, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE REFERENCE: PR/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 6  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-6

Query Match 93.0%; Score 66; DB 4; Length 24;  
Best Local Similarity 91.7%; Pred. No. 7.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
||| ||| ||| ||| |||  
Db 3 CFQWQENMRKVR 14

## RESULT 14

US-07-755-161A-10  
Sequence 10, Application US/07755161A  
Patent No. 5304633  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/755,161A  
FILING DATE: 19910905  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 21  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-10  
Query Match 93.0%; Score 66; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 7.5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQENMRKVR 12  
||| ||| ||| ||| |||  
Db 4 CFQWQENMRKVR 15

RESULT 15  
US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317984  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10  
; Query Match 93.0%; Score 66; DB 1; Length 25;  
; Best Local Similarity 91.7%; Pred. No. 7.5e-05;  
; Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
; QY 1 CFQWQNNRKVR 12  
; Db 4 CFQWQNNRKVR 15  
; Search completed: February 21, 2003, 07:50:37  
; Job time : 9.7 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-85  
Perfect score: 71  
Sequence: 1 CFQWQENMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	93.0	15	9	US-09-798-869-2
2	66	93.0	25	9	US-09-798-869-20
3	66	93.0	694	9	US-10-023-096-2
4	63	88.7	15	9	US-09-798-869-6
5	49	69.0	15	9	US-09-798-869-3
6	49	69.0	25	9	US-09-798-869-23
7	46	64.8	15	9	US-09-798-869-7
8	45	63.4	15	9	US-09-798-869-4
9	45	63.4	25	9	US-09-798-869-22
10	42	59.2	15	9	US-09-798-869-8
11	42	59.2	15	9	US-09-798-869-29
12	42	59.2	15	9	US-09-798-869-30
13	40	56.3	239	10	US-09-864-761-37353
14	37	52.1	747	9	US-10-066-500-58
15	37	52.1	747	9	US-10-002-796-58
16	37	52.1	747	9	US-10-066-273-58
17	37	52.1	747	9	US-10-066-494-58
18	37	52.1	817	10	US-09-976-165-28
19	37	52.1	829	10	US-09-976-165-31

20	37	52.1	2273	10	US-09-995-542-12	Sequence 12, Appl
21	37	52.1	2310	10	US-09-995-542-10	Sequence 10, Appl
22	36	50.7	33	10	US-09-864-761-40025	Sequence 40025, A
23	35	49.3	15	9	US-09-798-869-5	Sequence 5, Appl
24	35	49.3	119	10	US-09-205-658-244	Sequence 244, App
25	35	49.3	302	10	US-09-945-301-11	Sequence 11, Appl
26	35	49.3	509	10	US-09-879-957-194	Sequence 194, App
27	34.5	48.6	70	10	US-09-864-761-47165	Sequence 47165, A
28	34	47.9	62	10	US-09-815-242-12129	Sequence 12129, A
29	34	47.9	62	10	US-09-815-242-13026	Sequence 13026, A
30	34	47.9	321	10	US-09-771-730-43	Sequence 43, Appl
31	34	47.9	943	9	US-09-738-626-5199	Sequence 5199, Ap
32	34	47.9	1057	9	US-10-135-322-5	Sequence 5, Appl
33	34	47.9	1057	9	US-09-918-508-6	Sequence 6, Appl
34	34	47.9	1057	9	US-09-798-869-26	Sequence 26, Appl
35	33	46.5	13	9	US-09-798-869-25	Sequence 25, Appl
36	33	46.5	14	9	US-09-798-869-1	Sequence 1, Appl
37	33	46.5	15	9	US-09-798-869-10	Sequence 10, Appl
38	33	46.5	15	9	US-09-798-869-28	Sequence 28, Appl
39	33	46.5	15	9	US-09-798-869-24	Sequence 24, Appl
40	33	46.5	18	9	US-09-798-869-21	Sequence 21, Appl
41	33	46.5	25	9	US-09-909-652-4	Sequence 4, Appl
42	33	46.5	25	10	US-09-030-619-205	Sequence 205, App
43	33	46.5	25	10	US-09-917-340-16	Sequence 16, Appl
44	33	46.5	25	10	US-09-864-761-46393	Sequence 46393, A
45	33	46.5	40	10		

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 93.0%; Score 66; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 6.5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
Db 3 CFQWQENMRKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON

APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 93.0%; Score 66; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
| | | | |  
DB 3 CFQWQENMRKVR 14

RESULT 3  
US-10-023-096-2  
Sequence 2, Application US/10023096  
Patent No. US20020160941A1  
GENERAL INFORMATION:  
APPLICANT: Kruzel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
TITLE OF INVENTION: Lactoferrin  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,096  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 93.0%; Score 66; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0028;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
| | | | |  
DB 22 CFQWQENMRKVR 33

RESULT 4  
US-09-798-869-6  
Sequence 6, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE: OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 88.7%; Score 63; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.0002;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKVR 12  
| | | | |  
DB 3 CFQWQENMRKVR 14

RESULT 5  
US-09-798-869-3  
Sequence 3, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15  
TYPE: PRT  
ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 69.0%; Score 49; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.037;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQRRMRKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 69.0%; Score 49; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.062;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQRRMRKL 13

## RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 64.8%; Score 46; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.11;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQRRMRKL 13

## RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 63.4%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.17;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQENMRKV 13

## RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 63.4%; Score 45; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.27;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQENMRKV 13



RESULT 10  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.51;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
| : || ||||  
Db 3 CLRQWQWRMKKL 13

RESULT 11  
US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.51;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
| : || ||||  
Db 3 CLRQWQWRMKKL 13

RESULT 12  
US-09-798-869-30

; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.51;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
| : || ||||  
Db 3 CLRQWQWRMKKL 13

RESULT 13  
US-09-864-761-37353  
; Sequence 37353, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aesomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR FILING DATE: 2001-01-30
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37353
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000114.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.83
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.84
; OTHER INFORMATION: EST HUMAN HIT: A1678113.1, EVALUATE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P36677, EVALUATE 5.00e-03
; US-09-864-761-37353

Query Match 56.3% Score 40; DB 10; Length 239;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENNRKV 11
| :|||:|
Db 164 CEQWENIRKL 174

RESULT 14
US-10-066-500-58
; Sequence 58, Application US/10066500
; Patent No. US20020177165A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Botstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Raton
; APPLICANT: Napoleone Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3130R1C7

; CURRENT APPLICATION NUMBER: US/10/066,500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106032
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/109304
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/139695
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/145070
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/149396
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 08/918874
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 08/933821
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 08/960507
```

; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 09/114844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: 09/136801  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136804  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/136828  
; PRIOR FILING DATE: 1998-08-19  
; PRIOR APPLICATION NUMBER: 09/158342  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: 09/180997  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 09/202088  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: 09/254311  
; PRIOR FILING DATE: 1999-03-03  
; PRIOR APPLICATION NUMBER: 09/254460  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: 09/254465  
; PRIOR FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: 09/284663  
; PRIOR FILING DATE: 1999-04-15  
; PRIOR APPLICATION NUMBER: 09/332928  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/332929  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333075  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/333077  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380139  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/403296  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/403297  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/423741  
; PRIOR FILING DATE: 1999-11-10  
; PRIOR APPLICATION NUMBER: 09/423844  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 09/522342  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 09/548815  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: 09/664610  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/665350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: 09/709238  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/767609  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 09/802706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 09/808689  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/866028  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 09/870574  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 09/872035  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/886342  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: PCT/US98/14552  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US98/18824  
; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/US98/19093  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: PCT/US98/19330  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: PCT/US98/19437  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/24855  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: PCT/US98/25108  
; PRIOR FILING DATE: 1998-12-01  
; PRIOR APPLICATION NUMBER: PCT/US98/25190  
; PRIOR FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: PCT/US99/20111  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 52.1%; Score 37; DB 9; Length 747;  
Best Local Similarity 45.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11  
| : | | | : :  
Db 311 CVRWQINSRRI 321

## RESULT 15

US-10-002-796-58  
; Sequence 58, Application US/10002796  
; Publication No. US20030032057A1  
; GENERAL INFORMATION:  
; APPLICANT: Avi J. Ashkenazi  
; APPLICANT: Kevin P. Baker  
; APPLICANT: David A. Botstein  
; APPLICANT: Luc Desnoyers  
; APPLICANT: Dan L. Eaton  
; APPLICANT: Napoleone Ferrara  
; APPLICANT: Sherman Fong  
; APPLICANT: Wei-Qiang Gao  
; APPLICANT: Hanspeter Gerber  
; APPLICANT: Mary E. Gerritsen  
; APPLICANT: Audrey Goddard  
; APPLICANT: Paul J. Godowski  
; APPLICANT: Austin L. Gurney  
; APPLICANT: Ivar J. Kljavin  
; APPLICANT: Jennie P. Mather  
; APPLICANT: Mary A. Napier  
; APPLICANT: James Pan  
; APPLICANT: Nicholas F. Paoni  
; APPLICANT: Margaret Ann Roy  
; APPLICANT: Timothy A. Stewart  
; APPLICANT: Daniel Tumas  
; APPLICANT: Colin K. Watanabe  
; APPLICANT: P. Mickey Williams  
; APPLICANT: William I. Wood  
; APPLICANT: Zemin Zang  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3130R1C1  
; CURRENT APPLICATION NUMBER: US/10/002,796  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263

;; PRIOR FILING DATE: 1997-09-18  
;; PRIOR APPLICATION NUMBER: 60/059588  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/062285  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/062816  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/063082  
;; PRIOR FILING DATE: 1997-10-31  
;; PRIOR APPLICATION NUMBER: 60/063329  
;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063733  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066840  
;; PRIOR FILING DATE: 1997-11-25  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/095998  
;; PRIOR FILING DATE: 1998-08-10  
;; PRIOR APPLICATION NUMBER: 60/097000  
;; PRIOR FILING DATE: 1998-08-18  
;; PRIOR APPLICATION NUMBER: 60/099601  
;; PRIOR FILING DATE: 1998-09-09  
;; PRIOR APPLICATION NUMBER: 60/099803  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/099811  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/099812  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/100858  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/101922  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/106032  
;; PRIOR FILING DATE: 1998-10-28  
;; PRIOR APPLICATION NUMBER: 60/109304  
;; PRIOR FILING DATE: 1998-11-20  
;; PRIOR APPLICATION NUMBER: 60/125778  
;; PRIOR FILING DATE: 1999-03-23  
;; PRIOR APPLICATION NUMBER: 60/139695  
;; PRIOR FILING DATE: 1999-06-15  
;; PRIOR APPLICATION NUMBER: 60/145070  
;; PRIOR FILING DATE: 1999-07-20  
;; PRIOR APPLICATION NUMBER: 60/145698  
;; PRIOR FILING DATE: 1999-07-26  
;; PRIOR APPLICATION NUMBER: 60/149396  
;; PRIOR FILING DATE: 1999-08-17  
;; PRIOR APPLICATION NUMBER: 60/169495  
;; PRIOR FILING DATE: 1999-12-07  
;; PRIOR APPLICATION NUMBER: 08/018874  
;; PRIOR FILING DATE: 1997-08-26  
;; PRIOR APPLICATION NUMBER: 08/933821  
;; PRIOR FILING DATE: 1997-09-19  
;; PRIOR APPLICATION NUMBER: 08/960507  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 09/114844  
;; PRIOR FILING DATE: 1998-07-14  
;; PRIOR APPLICATION NUMBER: 09/136801  
;; PRIOR FILING DATE: 1998-08-19  
;; PRIOR APPLICATION NUMBER: 09/136804  
;; PRIOR FILING DATE: 1998-08-19  
;; PRIOR APPLICATION NUMBER: 09/136828  
;; PRIOR FILING DATE: 1998-08-19

;; PRIOR APPLICATION NUMBER: 09/158342  
;; PRIOR FILING DATE: 1998-09-21  
;; PRIOR APPLICATION NUMBER: 09/180997  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 09/202088  
;; PRIOR FILING DATE: 1998-12-08  
;; PRIOR APPLICATION NUMBER: 09/254311  
;; PRIOR FILING DATE: 1999-03-03  
;; PRIOR APPLICATION NUMBER: 09/254460  
;; PRIOR FILING DATE: 1999-03-09  
;; PRIOR APPLICATION NUMBER: 09/254465  
;; PRIOR FILING DATE: 1999-03-05  
;; PRIOR APPLICATION NUMBER: 09/284663  
;; PRIOR FILING DATE: 1999-04-15  
;; PRIOR APPLICATION NUMBER: 09/332928  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/332929  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/333075  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/333077  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/380137  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/380138  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/380139  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/403296  
;; PRIOR FILING DATE: 1999-10-18  
;; PRIOR APPLICATION NUMBER: 09/403297  
;; PRIOR FILING DATE: 1999-10-18  
;; PRIOR APPLICATION NUMBER: 09/423741  
;; PRIOR FILING DATE: 1999-11-10  
;; PRIOR APPLICATION NUMBER: 09/423844  
;; PRIOR FILING DATE: 1999-11-12  
;; PRIOR APPLICATION NUMBER: 09/522342  
;; PRIOR FILING DATE: 2000-03-09  
;; PRIOR APPLICATION NUMBER: 09/548815  
;; PRIOR FILING DATE: 2000-04-13  
;; PRIOR APPLICATION NUMBER: 09/664610  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: 09/665350  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: 09/709238  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 09/767609  
;; PRIOR FILING DATE: 2001-01-22  
;; PRIOR APPLICATION NUMBER: 09/802706  
;; PRIOR FILING DATE: 2001-03-09  
;; PRIOR APPLICATION NUMBER: 09/808689  
;; PRIOR FILING DATE: 2001-03-14  
;; PRIOR APPLICATION NUMBER: 09/866028  
;; PRIOR FILING DATE: 2001-05-25  
;; PRIOR APPLICATION NUMBER: 09/870574  
;; PRIOR FILING DATE: 2001-05-30  
;; PRIOR APPLICATION NUMBER: 09/872035  
;; PRIOR FILING DATE: 2001-06-01  
;; PRIOR APPLICATION NUMBER: 09/886342  
;; PRIOR FILING DATE: 2001-06-19  
;; PRIOR APPLICATION NUMBER: PCT/US98/14552  
;; PRIOR FILING DATE: 1998-07-14  
;; PRIOR APPLICATION NUMBER: PCT/US98/18824  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: PCT/US98/19093  
;; PRIOR FILING DATE: 1998-09-14  
;; PRIOR APPLICATION NUMBER: PCT/US98/19330  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: PCT/US98/19437  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: PCT/US98/24855  
;; PRIOR FILING DATE: 1998-11-20  
;; PRIOR APPLICATION NUMBER: PCT/US98/25108

; PRIOR FILING DATE: 1998-12-01  
; PRIOR APPLICATION NUMBER: PCT/US98/25190  
; PRIOR FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: PCT/US99/20111  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/28301

Query Match 52.1%; Score 37; DB 9; Length 747;  
Best Local Similarity 45.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWQENMKV 11  
| : | | | : |  
Db 311 CWRWQINGRI 321

Search completed: February 21, 2003, 08:08:09  
Job time : 11.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 Seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107B-85

Perfect score: 71

Sequence: 1 CFQWQNMRRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:\*

1: piri:\*

2: piri:\*

3: piri:\*

4: piri:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	93.0	711	1 TFHUL	lactotransferrin p
2	50	70.4	33	2 S52107	lactoferrin - shee
3	49	69.0	708	2 JC2323	lactoferrin - goat
4	45	63.4	707	1 A28438	lactoferrin precu
5	45	63.4	4568	2 T08030	dyein beta heavy
6	42	59.2	289	2 G86403	33.3K hypothetical
7	41	57.7	275	2 T22597	hypothetical prote
8	41	57.7	511	2 AB0858	hypothetical prote
9	40	56.3	415	2 C71467	probable tyrosine
10	40	56.3	431	2 S50977	hypothetical prote
11	39	54.9	267	2 S77802	hypothetical prote
12	39	54.9	660	2 F98169	serine proteinase
13	39	54.9	660	2 AB3118	serine proteinase
14	39	54.9	681	2 T19429	hypothetical prote
15	39	54.9	966	1 P1BVBB	RNA 1a protein - b
16	39	54.9	1174	2 C97686	pyruvate carboxyla
17	39	54.9	1174	2 AE2911	pyruvate carboxyla
18	39	54.9	1208	2 T00362	hypothetical prote
19	39	54.9	1432	2 B85431	trichovyalin like
20	38	53.5	206	2 H97451	pyridoxamine 5'-ph
21	38	53.5	206	2 AB2570	pyridoxamine 5'-ph
22	38	53.5	208	2 AG3441	probable pyridoxam
23	38	53.5	282	2 F90580	hypothetical prote
24	38	53.5	357	2 T01607	probable triacylgl
25	38	53.5	487	2 A82835	anthranilate synth
26	38	53.5	572	1 DERTMX	malate dehydrogena
27	38	53.5	572	2 S44415	malate dehydrogena
28	38	53.5	584	2 C84325	hypothetical prote
29	38	53.5	602	2 E64464	hypothetical prote

## ALIGNMENTS

### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N;Alternate names: lactoferrin

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C;Accession: G01394; S1228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R;Cho, Y.

submitted to the EMBL Data Library, March 1994

A;Reference number: G06820

A;Accession: G01394

A;Status: preliminary; translated from GB/EMBL/DBD

A;Molecule type: mRNA

A;Residues: 1-711 <CHO>

A;Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R;Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A;Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A;Reference number: S11228; MUID:90384839; PMID:2402455

A;Accession: S11228

A;Molecule type: mRNA

A;Residues: 1-148,'T',150-422,'C',424-711 <REY>

A;Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R;Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A;Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A;Reference number: A45401; MUID:93125571; PMID:1480183

A;Accession: A45401

A;Molecule type: DNA

A;Residues: 1-15 <TEN>

A;Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A;Experimental source: placenta

A;Note: sequence extracted from NCBI backbone (NCBIP:122202)

R;Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A;Title: Nucleotide sequence of human lactoferrin cDNA.

A;Reference number: S10324; MUID:90326549; PMID:2374734

A;Accession: S10324

A;Molecule type: mRNA

A;Residues: 3-711 <POW>

A;Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R;Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A;Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A;Reference number: S15853; MUID:91264786; PMID:2049066

A;Accession: S15853

A;Status: nucleic acid

A;Molecule type: mRNA

A;Residues: 20-31 <ST1>

A;Accession: S20841

A;Molecule type: protein

A;Residues: 20-28,'X',30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:q186815; PIDN:AAA86665.1; PID:9386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SMKPVN' <PAN>  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LIF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3Q21-3Q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-71/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status e  
 F;157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

Query Match 93.0%; Score 66; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0016;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQENNRKV 12  
 Db 39 CFQWQENNRKV 50  
 ||||| |||||  
 ||||| |||||

RESULT 2  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet a  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 70.4%; Score 50; DB 2; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.04;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQENNRKV 11  
 Db 19 CYQWQENNRKL 29  
 ||||| |||||  
 ||||| |||||

RESULT 3  
 JC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JC2323  
 R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locu  
 A;Reference number: JC2323; MUID:94380047; PMID:8093048  
 A;Accession: JC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.0%; Score 49; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 1.4;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQENNRKV 11  
 Db 38 CYQWQENNRKL 48  
 ||||| |||||  
 ||||| |||||

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secr  
 A;Reference number: A92596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.4%; Score 45; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 7;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWQENNRKV 11  
 Db 37 CLRQWQENNRKV 47  
 ||||| |||||  
 ||||| |||||

T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T22597  
 R:Dobson, R.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19587  
 A:Accession: T22597  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-275 <WIL>  
 A:Cross-references: EMBL:Z81089; PIDN:CA031137.1; GSPDB:GNO0028; CESP:F53H4.4  
 A:Experimental source: clone F53H4  
 C:Genetics:  
 A:Gene: CESP:F53H4.4  
 A:Map position: X  
 A:Introns: 67/1; 153/1  
 C:Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 57.7%; Score 41; DB 2; Length 275;  
 Best Local Similarity 63.6%; Pred. No. 13;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FQWQENMRKVR 12  
 ||||| :|||  
 Db 262 FQWQENMRKVR 272

RESULT 8  
 AB0858  
 hypothetical protein STY3070 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C:Species: Salmonella enterica subsp. enterica serovar Typhimurium  
 C>Note: this species has also been called Salmonella typhi  
 C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AB0858  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A:Reference number: AB0858; PMID:11677608  
 A:Accession: AB0858  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-511 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:gl6504016; GSPDB:GNO0176  
 C:Genetics:  
 A:Gene: STY3070

Query Match 57.7%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 CFQWQENMRKVR 12  
 ||||| :|||  
 Db 350 CFQWQENMRKVR 361

RESULT 9  
 C71467  
 probable tyrosine transport - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: C71467  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.; Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71570; MUID:99000809; PMID:9784136  
 A:Accession: C71467  
 A>Status: preliminary  
 A:Molecule type: DNA

Query Match 59.2%; Score 42; DB 2; Length 289;  
 Best Local Similarity 58.3%; Pred. No. 9.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWQENMRKVR 12  
 ||||| :|||  
 Db 8 CFTWEYARHVR 19

RESULT 7  
 A:Map position: 1

Query Match 59.2%; Score 42; DB 2; Length 289;  
 Best Local Similarity 58.3%; Pred. No. 9.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWQENMRKVR 12  
 ||||| :|||  
 Db 8 CFTWEYARHVR 19

RESULT 7  
 A:Map position: 1

Query Match 63.4%; Score 45; DB 2; Length 4568;  
 Best Local Similarity 50.0%; Pred. No. 48;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFQWQENMRKVR 12  
 ||||| :|||  
 Db 1852 CFQWQENMRKVR 1863

RESULT 6  
 G86403  
 33.3K hypothetical protein F28L5.13 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
 C:Accession: G86403  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Hughes, N.F.; Huizar, L.; Hunter, J.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, B.; Weller, D.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, B.; Weller, D.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G86403  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-289 <STO>  
 A:Cross-references: GB:AE005172; NID:gl0998925; PIDN:AAG26065.1; GSPDB:GNO0141  
 C:Genetics:  
 A:Map position: 1

Query Match 59.2%; Score 42; DB 2; Length 289;  
 Best Local Similarity 58.3%; Pred. No. 9.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWQENMRKVR 12  
 ||||| :|||  
 Db 8 CFTWEYARHVR 19

RESULT 7  
 A:Map position: 1



A:Residues: 1-415 <ARN>  
A:Cross-references: GB:AE001354; GB:AE001273; NID:G3329280; PIDN:AA068414.1; PID:G3329280  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: tyrP\_1  
C:Superfamily: tyrosine-specific transport protein

Query Match 56.3%; Score 40; DB 2; Length 415;  
Best Local Similarity 70.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWQENMRKV 11  
||| |||  
Db 317 FQWDEKKRV 326

RESULT 10  
S50977  
hypothetical protein YDL005c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein D2930; hypothetical protein YD8119.02c  
C:Species: Saccharomyces cerevisiae  
C:Date: 11-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 19-Apr-2002  
C:Accession: S50977; S52514; S67537  
R:Murphy, L.; Richards, C.; Gentiles, S.; Harris, D.  
submitted to the EMBL Data Library, January 1995  
A:Reference number: S50976  
A:Accession: S50977  
A:Molecule type: DNA  
A:Residues: 1-431 <MUR>  
A:Cross-references: EMBL:Z48008; NID:G642799; PIDN:CAA88056.1; PID:G642801  
R:Andre, B.; Vissers, S.; Urrestarazu, L.  
submitted to the EMBL Data Library, February 1995  
A:Description: The sequence of a 42 kb segment located on the left arm of chromosome IV  
A:Reference number: S52492  
A:Accession: S52514  
A:Molecule type: DNA  
A:Residues: 1-431 <AND>  
A:Cross-references: EMBL:Z48432; NID:G683669; PIDN:CAA88354.1; PID:G683692  
R:Urrestarazu, L.A.; Andre, B.; Vissers, S.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67535  
A:Accession: S67537  
A:Molecule type: DNA  
A:Residues: 1-431 <URR>  
A:Cross-references: EMBL:Z74053; NID:G1430962; PIDN:CAA98561.1; PID:G1430962  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:MED2  
A:Cross-references: SGD:S0002163  
A:Map position: 4L

Query Match 56.3%; Score 40; DB 2; Length 431;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QENMRKV 12  
||| |||  
Db 122 QENMRKV 129

RESULT 11  
S77802  
hypothetical protein MC003 - Mycoplasma capricolum (fragment)  
C:Species: Mycoplasma capricolum  
C:Date: 09-Oct-1997 #sequence\_revision 31-Oct-1997 #text\_change 07-Dec-1999  
C:Accession: S77802  
R:Horik, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.; Mol. Microbiol. 16, 955-967, 1995  
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiology  
A:Reference number: S77739; MUID:96059641; PMID:7476192  
A:Accession: S77802  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-267 <BOR>  
A:Cross-references: EMBL:Z33006  
A:Experimental source: ATCC 27343  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Genetics:  
A:Genetic code: SGC3

Query Match 54.9%; Score 39; DB 2; Length 267;  
Best Local Similarity 58.3%; Pred. No. 28;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 12  
||| |||  
Db 158 CFQKKNMRQMR 169

RESULT 12  
E98169  
serine proteinase XF0267 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: E98169  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; PMID:11743194  
A:Accession: E98169  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-660 <KJR>  
A:Cross-references: GB:AE007870; PIDN:AAK88879.1; PID:G15158646; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L\_612  
A:Map position: linear chromosome

Query Match 54.9%; Score 39; DB 2; Length 660;  
Best Local Similarity 87.5%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WOENMRKV 11  
||| |||  
Db 126 WOENMRKV 133

RESULT 13  
AB3118  
serine proteinase Atu4566 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AB3118  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S. Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AB3118  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-660 <KJR>  
A:Cross-references: GB:AB008689; PIDN:AA45360.1; PID:G17743054; GSPDB:GN00187  
C:Genetics:  
A:Gene: Atu4566  
A:Map position: linear chromosome

Query Match 54.9%; Score 39; DB 2; Length 660;  
Best Local Similarity 87.5%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WOENMRKV 11  
DB 126 WOENMRKV 133

## RESULT 14

T19429  
hypothetical protein C24H11.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19429  
R:Lloyd, C.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19123  
A:Accession: T19429  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-681 <WIL>  
A:Cross-references: EMBL:Z81475; PIDN:CAB03914.1; GSPDB:GN00021; CESP:C24H11.8  
C:Genetics:  
A:Gene: CESP:C24H11.8  
A:Map position: 3  
A:Introns: 18/3; 65/3; 108/1; 138/1; 177/3; 256/2; 303/2; 343/1; 409/3; 484/3; 532/2; 58

Query Match 54.9%; Score 39; DB 2; Length 681;  
Best Local Similarity 60.0%; Pred. NO. 74;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 WOENMRKV 12  
DB 192 RWQNRVRV 201

## RESULT 15

PIBVBS  
RNA la protein - broad bean mottle virus (strain Bawden)  
N:Contains: ATP-dependent helicase (EC 3.6.1.-); mRNA (guanine-N7-)-methyltransferase (E  
C:Species: broad bean mottle virus  
C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 19-Jan-2001  
C:Accession: A41699  
R:Dzianott, A.M.; Bujarski, J.J.  
Virology 185, 553-562, 1991  
A:Title: The nucleotide sequence and genome organization of the RNA-1 segment in two bro  
A:Reference number: A41699; MUID:92074218; PMID:1962437  
A:Accession: A41699  
A:Molecule type: genomic RNA  
A:Residues: 1-966 <DZI>  
A:Cross-references: GB:M65138; MID:g210662; PIDN:AAA42740.1; PID:g210663  
C:Superfamily: cucumber mosaic virus RNA 1 protein  
C:Keywords: hydrolase; methyltransferase; mRNA capping; nucleotide binding; P-loop; S-ad  
F:80-168/Domain: methyltransferase #status predicted <MTF>  
F:688-944/Domain: helicase #status predicted <HHG>  
F:690-697/Region: nucleotide-binding motif A (P-loop)

Query Match 54.9%; Score 39; DB 1; Length 966;  
Best Local Similarity 53.3%; Pred. NO. 1.1e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 CFQ----WOENMRKV 11  
DB 347 CFQNKDWTENMRV 361

Search completed: February 21, 2003, 07:47:57  
Job time : 10.65 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 Seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-85  
Perfect score: 71  
Sequence: 1 CFQWQENNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	66	93.0	711	1	TRFL_HUMAN	P02788 homo sapien
2	49	69.0	708	1	TRFL_CAMDR	Q9tum0 camelus dro
3	49	69.0	708	1	TRFL_CAPHI	Q29477 capra hircu
4	45	63.4	707	1	TRFL_MOUSE	P08071 mus musculu
5	45	63.4	4568	1	DYHB_CHLRE	Q39565 chlamydomon
6	41	57.7	695	1	TRFL_HORSE	O77811 equus cabal
7	39	54.9	267	1	V125_MYCCA	P3661 mycoplasma
8	39	54.9	966	1	VIA_BBMV	Q00020 broad bean
9	39	54.9	1137	1	PHYC_ORYSA	Q92w19 oryza sativ
10	38	53.5	62	1	RL28_THETN	Q8r9u1 thermocanaer
11	38	53.5	333	1	CATJ_MOUSE	Q9r014 mus musculu
12	38	53.5	572	1	MAOX_HUMAN	P48163 homo sapien
13	38	53.5	572	1	MAOX_RAT	P13697 rattus norv
14	38	53.5	857	1	EF2_CHICK	Q90705 gallus gall
15	38	53.5	857	1	EF2_CRIGR	P09445 cricetus
16	38	53.5	857	1	EF2_HUMAN	P13639 homo sapien
17	38	53.5	857	1	EF2_MESAU	P05086 mesocricetu
18	38	53.5	857	1	EF2_MOUSE	P58252 mus musculu
19	38	53.5	857	1	EF2_RAT	P05197 rattus norv
20	38	53.5	892	1	RA16_SCHPO	P36617 schizosacch
21	37	52.1	172	1	V708_METJA	Q58118 methanococc
22	37	52.1	217	1	YNJD_ECOLI	P76909 escherichia
23	37	52.1	612	1	RPSD_BUCAI	P57163 buchnera ap
24	37	52.1	2273	1	ABCR_HUMAN	P78363 homo sapien
25	37	52.1	2483	1	COA2_HUMAN	Q00763 homo sapien
26	36	50.7	238	1	YBM9_SCHPO	Q10333 schizosacch
27	36	50.7	306	1	BUB2_YEAST	P26448 saccharomyc
28	36	50.7	317	1	MOCA_RHIME	P49307 rhizobium m
29	36	50.7	349	1	F16P_ANASP	P48991 anabaena sp
30	36	50.7	358	1	DDL_ENTHR	Q47827 enterococcu
31	36	50.7	396	1	YUIB_BACSU	Q34374 bacillus su
32	36	50.7	455	1	YAYL_CAEEL	Q19910 caenorhabdi
33	36	50.7	482	1	U2R2_HUMAN	Q15695 homo sapien

34 36 50.7 749 1 VP4\_ROTGA Q04916 rotavirus (

35 36 50.7 851 1 EF2\_CAEEL P29691 caenorhabdi

36 36 50.7 961 1 VIA\_BMV P03588 brome mosai

37 36 50.7 962 1 YBX7\_SCHPO P10201 schizosacch

38 35.5 50.0 282 1 LFXC\_AQUAB O67648 aquifex asc

39 35 49.3 57 1 YE56\_ARCFU O28816 archaeoglob

40 35 49.3 343 1 CATQ\_RAT Q9qz3 rattus norv

41 35 49.3 355 1 MURG\_NEIMA O9fz27 neisseria m

42 35 49.3 355 1 MURG\_NEIMB Q9k0y2 neisseria m

43 35 49.3 430 1 SYH\_CHLPN Q9z7p1 chlamydia p

44 35 49.3 475 1 YEBA\_HAEIN P44693 haemophilus

45 35 49.3 479 1 U2R1\_HUMAN Q15695 homo sapien

## ALIGNMENTS

RESULT 1

TRFL\_HUMAN

ID TRFL\_HUMAN STANDARD; PRT: 711 AA.

AC P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;

AC Q96KZ5;

DT 21-JUL-1996 (Rel. 01, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lactoferrin precursor (Lactoferrin) [Contains: Lactoferrin A;

DE Lactoferrin B; Lactoferrin C].

GN LTF OR LF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RX MEDLINE=90384839; PubMed=2402455;

RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;

RT "Complete nucleotide sequence of human mammary gland lactoferrin.";

RL Nucleic Acids Res. 18:5288-5288(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RX

RA Cho Y.Y.;

RT Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.

RN [3]

RP SEQUENCE FROM N.A.

RA Connely O.M.;

RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RX

RA Liang Q., Jimenez-Flores R., Richardson T.;

RT "Molecular cloning and sequence analysis of human lactoferrin.";

RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RA Wei X., Han J., Rado T.A.;

RT "Human neutrophil lactoferrin coding and 5' flanking region DNA

sequences.";

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RA Cheng H., Chen X., Huan L.;

RT "cDNA cloning and sequence analysis of human lactoferrin.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [7]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Oden J.E.,  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RT Legrand D., Spik G., Montreuil J., Jolles P.;  
RL "Human lactoferrin: amino acid sequence and structural  
RT comparisons with other transferrins";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RN Nian M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Guo J., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RN Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253-->-methionine mutant";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;

RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier P., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene";  
RL Mol. Vision 4:31-32(1998).  
RN [21]  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC  
CC -I- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC  
CC -I- SUBUNIT: MONOMER.  
CC  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC  
CC -I- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC  
CC -I- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83202; AAA59511.1; -  
CC EMBL; M83203; AAA58656.1; -  
CC EMBL; M18642; AAA86665.1; -  
CC EMBL; AF332168; AAC48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC  
CC PDB; 1LCF; 31-AUG-94.  
CC PDB; 1LCT; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGG; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.

Query Match 93.0%; Score 66; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00037;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12  
 ||||| |||||  
 DB 39 CFQWQENMKVR 50

## RESULT 2

TRFL CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9MZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]\_TaxID=9838;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AJ131674; CAB53387.1; -.  
 DR EMBL; AF165879; AAF82241.1; -.  
 DR HSP; O77811; IBIX.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF04005; Transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.

FT SIGNAL. 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 ANION 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 69.0%; Score 49; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.139;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12  
 ||||| |||||  
 DB 38 CAQWQENMKVR 49

## RESULT 3

TRFL CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Mocart M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 RT relevant locus to bovine U12 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332 (1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.



[illegible]

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AJ010930; CAA09407.1; --
PDB; 1B1X; 02-DEC-98.
PDB; 1B7U; 02-FEB-99.
PDB; 1B7Z; 02-FEB-99.
InterPro; IPR001156; Transferrin.
Pfam; PF00405; transferrin; 2.
PRINTS; PR00422; TRANSFERRIN.
SMART; SM00094; TR_FER; 2.
PROSITE; PS00205; TRANSFERRIN 1; 2.
PROSITE; PS00206; TRANSFERRIN 2; 2.
PROSITE; PS00207; TRANSFERRIN 3; 1.
Signal; 3D-structure. 1
NON_TER 1
CHAIN <1 6 LACTOTRANSFERRIN.
REPEAT 7 695
REPEAT 7 350 1.
DISULFID 15 51
DISULFID 25 42
DISULFID 121 204
DISULFID 163 179
DISULFID 166 189
DISULFID 176 187
DISULFID 237 251
DISULFID 354 386
DISULFID 364 377
DISULFID 411 690
DISULFID 431 653
DISULFID 463 538
DISULFID 487 681
DISULFID 497 511
DISULFID 508 521
DISULFID 579 593
DISULFID 631 636
METAL 66 66
METAL 98 98
METAL 198 198
METAL 259 259
METAL 401 401
METAL 439 439
METAL 532 532
METAL 601 601
METAL 127 127
BINDING 469 469
BINDING 143 143
CARBOHYD 287 287
CARBOHYD 482 482
SEQUENCE 695 AA; 07BB84D50E1B165D CRC64;

Query Match 57.7%; Score 41; DB 1; Length 695;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12
Db 25 CAXFQNNMKVR 36

RESULT 7
Y125 MYCCA STANDARD; PRT; 267 AA.
AC P53661;
DT 01-OCT-1996 (Rel. 34, Created)

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z33006; CAA83689.1; --
InterPro; IPR001454; Hlgmase/hydr-lase.
InterPro; IPR000150; Hypothet_cof.
Pfam; PF00702; Hydrolase; 1.
PROSITE; PS01228; COF 1; 1.
PROSITE; PS01229; COF_2; 1.
Hypothetical protein.
NON_TER 267 267
SEQUENCE 267 AA; D5912DDB5839A8451 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 267;
Best Local Similarity 58.3%; Pred. No. 8.4;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12
Db 158 CFGKKNNQMR 169

RESULT 8
VIA_BMW STANDARD; PRT; 966 AA.
AC Q00020;
QY 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1A protein [includes: Helicase; Methyltransferase].
OC Broad bean mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Bromovirus.
NCBI_TaxID=12301;
OX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bawden;
RX MEDLINE=92074218; PubMed=1962437;
RA Dzianott A.M., Bujarski J.J.;
RT "The nucleotide sequence and genome organization of the RNA-1 segment
in two bromoviruses: broad bean mottle virus and cowpea chlorotic
mottle virus.";
RL Virology 185:553-562(1991).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.
CC CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE
CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA
CC CAPPING.
CC -!- SIMILARITY: TO 1A PROTEIN FROM CCMV, CMV, PSV AND TAV.
-----

```

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343 / KID;
RX MEDLINE=96059641; PubMed=7476192;
RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,
RA Dolan M., Gilbert W., Gillevet P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
its physiology.";
RL Mol. Microbiol. 16:955-967(1995).
CC -!- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E.COLI) / YCSE/YXEH
CC (B.SUBTILIS) FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z33006; CAA83689.1; --
InterPro; IPR001454; Hlgmase/hydr-lase.
InterPro; IPR000150; Hypothet_cof.
Pfam; PF00702; Hydrolase; 1.
PROSITE; PS01228; COF 1; 1.
PROSITE; PS01229; COF_2; 1.
Hypothetical protein.
NON_TER 267 267
SEQUENCE 267 AA; D5912DDB5839A8451 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 267;
Best Local Similarity 58.3%; Pred. No. 8.4;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMKVR 12
Db 158 CFGKKNNQMR 169

RESULT 8
VIA_BMW STANDARD; PRT; 966 AA.
AC Q00020;
QY 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1A protein [includes: Helicase; Methyltransferase].
OC Broad bean mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Bromovirus.
NCBI_TaxID=12301;
OX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bawden;
RX MEDLINE=92074218; PubMed=1962437;
RA Dzianott A.M., Bujarski J.J.;
RT "The nucleotide sequence and genome organization of the RNA-1 segment
in two bromoviruses: broad bean mottle virus and cowpea chlorotic
mottle virus.";
RL Virology 185:553-562(1991).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.
CC CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE
CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA
CC CAPPING.
CC -!- SIMILARITY: TO 1A PROTEIN FROM CCMV, CMV, PSV AND TAV.
-----

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M65138; AAA4740.1; -;  
 CC DR PIR; A41699; P1BVB.1;  
 CC DR InterPro; IPR002588; V\_methyltransf.  
 CC DR InterPro; IPR000606; Viral\_helicase1.  
 CC DR Pfam; PF01443; Viral\_helicase1; 1.  
 CC DR Pfam; PF01660; Vmethyltransf; 1.  
 CC DR Helicase; ATP-binding; Transferrase; Methyltransferase.  
 CC KW NP BIND 690 697 ATP (POTENTIAL)  
 CC FT SEQUENCE 966 AA; 109621 MW; DF592681D7231C8D CRC64;  
 CC SQ

Query Match 54.9%; Score 39; DB 1; Length 966;  
 Best Local Similarity 53.3%; Pred. No. 33;  
 Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 CFQ----WQENRKV 11  
 ||: |||||  
 DB 347 CFKKNKDWTEENRSV 361

## RESULT 9

PHYC\_ORYSA STANDARD; PRT; 1137 AA.  
 AC Q9ZWI9; P93429;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytochrome C.  
 GN PHYC.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Tahir M., Kanegae H., Takano M.;  
 RT "Phytochrome C (PHYC) gene in rice: isolation and characterization of  
 RT a complete coding sequence.";  
 RL (In) Plant Gene Register PGR98-210.  
 RN [2]  
 RP SEQUENCE OF 275-378 FROM N.A.  
 RX MEDLINE=97019052; PubMed=8865668;  
 RA Mathews S., Sharrock R.A.;  
 RT "The phytochrome gene family in grasses (Poaceae): a phylogeny and  
 RT evidence that grasses have a subset of the loci found in dicot  
 RT angiosperms.";  
 RL Mol. Biol. Evol. 13:1141-1150(1996).

CC -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT  
 CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS  
 CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT  
 CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN  
 CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS  
 CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE  
 CC RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
 CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RUBULOSE-  
 CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,  
 CC PROTOCHLOROPHYLLIDE REDUCTASE, RNMA, ETC. IT ALSO CONTROLS  
 CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.  
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -!- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.  
 CC -!- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
 CC -!- SIMILARITY: CONTAINS 2 PAS (PFR-ARNT-SIM) DIMERIZATION DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB018442; BAA74448.1; -;  
 CC EMBL; U61207; AAB41996.1; -;  
 CC DR InterPro; IPR003594; ATPbind\_ATPase.  
 CC DR InterPro; IPR003018; GAF.  
 CC DR InterPro; IPR004359; HIS\_KIN\_sig.  
 CC DR InterPro; IPR003661; His\_kinA.  
 CC DR InterPro; IPR001610; PAC.  
 CC DR InterPro; IPR000014; PAS domain.  
 CC DR InterPro; IPR001294; Phytochrome.  
 CC DR Pfam; PF00360; phytochrome; 1.  
 CC DR Pfam; PF00512; signal; 1.  
 CC DR Pfam; PF00989; PAS; 2.  
 CC DR Pfam; PF01590; GAF; 1.  
 CC DR Pfam; PF02518; HATPase\_C; 1.  
 CC DR PRINTS; PR01033; PHYTOCHROME.  
 CC DR SMART; SM00065; GAF; 1.  
 CC DR SMART; SM00387; HATPase\_C; 1.  
 CC DR SMART; SM00388; HSKA; 1.  
 CC DR SMART; SM00086; PAC; 1.  
 CC DR SMART; SM00091; PAS; 2.  
 CC DR TIGRFAMs; TIGR00229; sensory\_box; 2.  
 CC DR PROSITE; PS01109; HIS\_KIN; 1.  
 CC DR PROSITE; PS01112; PAS; 2.  
 CC DR PROSITE; PS00245; PHYTOCHROME\_1; 1.  
 CC DR PROSITE; PS00466; PHYTOCHROME\_2; 1.

KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;  
 Repeat; Multigene family.  
 FT DOMAIN 620 690 PAS 1.  
 FT DOMAIN 750 824 PAS 2.  
 FT DOMAIN 904 1124 HISTIDINE KINASE.  
 FT BINDING 322 322 CHROMOPHORE (BY SIMILARITY).  
 FT CONFLICT 279 279 F -> S (IN REF. 2).  
 FT CONFLICT 292 292 C -> S (IN REF. 2).  
 SQ SEQUENCE 1137 AA; 125982 MW; F2A520181CFE7B32 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 1137;  
 Best Local Similarity 45.5%; Pred. No. 39;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQENRKV 11  
 ||: |||||  
 DB 777 CLEWNEAQMKI 787

## RESULT 10

RL28\_THETN STANDARD; PRT; 62 AA.  
 AC Q8R9J1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L28.  
 GN RPMB OR TTE1495.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MBA / JCM 11007;  
 RX MEDLINE=21992816; PubMed=1197336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;

"A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AEO13107; ANM24713.1; -  
 CC Ribosomal protein; Complete proteome.  
 KW RIBOSOMAL PROTEIN; 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 62;  
 Best Local Similarity 60.0%; Pred. No. 2.7;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QWENNRKVR 12  
 :|:|:|:|:  
 DB 27 RAKPNRKVR 36  
 :|:|:|:|:  
 RESULT 11  
 CATJ MOUSE  
 ID CATJ\_MOUSE STANDARD; PRT; 333 AA.  
 AC Q9R0L4; Q9WV51;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cathepsin J precursor (EC 3.4.22.-) (Cathepsin P) (Catlp-p).  
 GN CTSP OR CTSP.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=99456833; PubMed=10526153;  
 RA Tislar K., Deussing J., Peters C.;  
 RT "Cathepsin J, a novel murine cysteine protease of the papain family  
 RT with a placenta-restricted expression.";  
 RL FEBS Lett. 459:299-304(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Placenta;  
 RA Sol-Church K., French J., Troeber D., Mason R.W.;  
 RT "Cloning of a mouse cysteine protease.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Lysosomal (Potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF136272; AAF13142.1; -  
 CC EMBL; AF156182; AAD41898.1; -  
 CC HSP; P07711; ICUL.  
 DR MEROPS; C01.038; -  
 DR MGD; MGI:1349426; Ctsj.  
 DR InterPro; IPR000668; Peptidase C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPAINE.

DR ProDom; PD000158; Peptidase C1; 1.  
 DR PROSITE; PS00139; THIOLESTERASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOLESTERASE\_HIS; 1.  
 DR PROSITE; PS00640; THIOLESTERASE\_ASN; 1.  
 KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.  
 FT SIGNAL 1 17  
 FT PROPEP 18 112  
 FT CHAIN 113 333  
 FT ACT\_SITE 137 137  
 FT ACT\_SITE 275 275  
 FT ACT\_SITE 299 299  
 FT DISULFID 134 177  
 FT DISULFID 168 210  
 FT DISULFID 268 321  
 FT CARBOHYD 71 71  
 FT CARBOHYD 216 216  
 FT CARBOHYD 220 220  
 FT CARBOHYD 267 267  
 FT CONFLICT 42 42  
 FT CONFLICT P -> PK (IN REF. 2).  
 SQ SEQUENCE 333 AA; 37147 MW; F9A8FF1D5A13B721 CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 333;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WOENNRKVR 12  
 :|:|:|:|:  
 DB 51 WEENNRWK 59  
 :|:|:|:|:  
 RESULT 12  
 MAOX HUMAN  
 ID MAOX\_HUMAN STANDARD; PRT; 572 AA.  
 AC P48153; Q16855; Q9BWX8; Q9UIY4; Q9HLW3;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).  
 GN MEL.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=White adipose tissue;  
 RX MEDLINE=94244767; PubMed=8187880;  
 RA Loeber G., Dworkin M.B., Infante A., Ahorn H.;  
 RT "Characterization of cytosolic malic enzyme in human tumor cells.";  
 RL FEBS Lett. 344:181-186(1994).  
 RN [2]  
 RP SEQUENCE OF 8-572 FROM N.A.  
 RX MEDLINE=96397682; PubMed=8804575;  
 RA Chou W.Y., Huang S.M., Chang G.G.;  
 RT "Nonidentity of the cDNA sequence of human breast cancer cell malic  
 RT enzyme to that from the normal human cell.";  
 RL J. Protein Chem. 15:273-279(1996).  
 RN [3]  
 RP SEQUENCE OF 27-572 FROM N.A.  
 RA Tracey A.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: (S)-malate + NADP(+) = pyruvate + CO(2) +  
 CC NADPH.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE MALIC ENZYMES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



```

CC      RIBOSOME.
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -|- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2
CC      (BY SIMILARITY).
CC      -|- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
CC      (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
CC      RIBOSYLATED BY DIPHTHERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A (BY
CC      SIMILARITY).
CC      -|- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC      EF-G/EF-2 SUBFAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; U46663; AAA87587.1; -.
CC      HSSP; P13551; 1FNM.
CC      InterPro; IPR000640; EFG_C.
CC      InterPro; IPR004161; EFTU_D2.
CC      InterPro; IPR000795; EF_GTPbind.
CC      InterPro; IPR005225; Small_GTP.
CC      Pfam; PF00009; GTP_EFTU; 1.
CC      Pfam; PF00679; EFG_C; 1.
CC      Pfam; PF03144; GTP_EFTU_D2; 1.
CC      TIGRfam; TIGR00231; small_GTP; 1.
CC      PROSITE; PS00301; EFATOR_GTP; 1.
CC      Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
CC      KW
CC      INIT MET 0 0
CC      NP_BIND 25 32 GTP (BY SIMILARITY).
CC      NP_BIND 103 107 GTP (BY SIMILARITY).
CC      NP_BIND 157 160 GTP (BY SIMILARITY).
CC      MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
CC      MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
CC      MOD_RES 714 714 DIPHTHAMIDE (BY SIMILARITY).
CC      SEQUENCE 857 AA; 95247 MW; 3680187581F519E6 CRC64;
CC
CC      Query Match 53.5%; Score 38; DB 1; Length 857;
CC      Best Local Similarity 47.4%; Pred. No. 44;
CC      Matches 9; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
CC
CC      QY 2 FQW-----QENMRKVR 12
CC      |||||
CC      Db 682 FQWATKEGULCEENMRGVR 700
CC
CC      RESULT 15
CC      EF2 CRIGR
CC      ID EF2 CRIGR STANDARD; PRT; 857 AA.
CC      AC P09445;
CC      DT 01-MAR-1989 (Rel. 10, Created)
CC      DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC      DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC      DE Elongation factor 2 (EF-2).
CC      GN EF2.
CC      OS Cricetus griseus (Chinese hamster).
CC      OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
CC      OC Cricetulus.
CC      OX NCBI_TaxID=10029;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RX MEDLINE=88198187; PubMed=2834376;
CC      RA Nakanishi T., Kohno K., Ishiura M., Ohashi H., Uchida T.;
CC      RT "Complete nucleotide sequence and characterization of the 5'-flanking
CC      region of mammalian elongation factor 2 gene.";
CC      RL J. Biol. Chem. 263:6384-6391(1988)
CC      -|- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC      OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC      RIBOSOME.

```

```

CC      -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -|- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
CC      (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
CC      RIBOSYLATED BY DIPHTHERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A.
CC      -|- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC      EF-G/EF-2 SUBFAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; J03200; AAA50386.1; -.
CC      PIR; A28660; A28660.
CC      HSSP; P13551; 1FNM.
CC      InterPro; IPR000640; EFG_C.
CC      InterPro; IPR004161; EFTU_D2.
CC      InterPro; IPR000795; EF_GTPbind.
CC      InterPro; IPR005225; Small_GTP.
CC      Pfam; PF00009; GTP_EFTU; 1.
CC      Pfam; PF00679; EFG_C; 1.
CC      Pfam; PF03144; GTP_EFTU_D2; 1.
CC      TIGRfam; TIGR00231; small_GTP; 1.
CC      PROSITE; PS00301; EFATOR_GTP; 1.
CC      Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
CC      KW
CC      INIT MET 0 0
CC      NP_BIND 25 32 GTP (BY SIMILARITY).
CC      NP_BIND 103 107 GTP (BY SIMILARITY).
CC      NP_BIND 157 160 GTP (BY SIMILARITY).
CC      MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
CC      MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
CC      MOD_RES 714 714 DIPHTHAMIDE.
CC      SEQUENCE 857 AA; 95264 MW; DCE0A55CFBCB886 CRC64;
CC
CC      Query Match 53.5%; Score 38; DB 1; Length 857;
CC      Best Local Similarity 47.4%; Pred. No. 44;
CC      Matches 9; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
CC
CC      QY 2 FQW-----QENMRKVR 12
CC      |||||
CC      Db 682 FQWATKEGALCEENMRGVR 700

```

Search completed: February 21, 2003, 07:28:00  
Job time : 5.6 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 Seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107B-85  
Perfect score: 71  
Sequence: 1 CFQWQNMRRKV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_rvirus: \*  
16: sp\_bacteriaph: \*  
17: sp\_archaea: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	85.9	711	4 Q8TCD2	Q8tcd2 homo sapien
2	57	80.3	38	4 Q9UCY5	Q9ucy5 homo sapien
3	50	70.4	33	6 Q9TR80	Q9tr80 ovis aries
4	42	59.2	289	10 Q9C6N2	Q9c6n2 arabidopsis
5	41	57.7	81	15 Q90863	Q90863 human immun
6	41	57.7	275	5 Q93780	Q93780 caenorhabdi
7	41	57.7	332	11 Q9DAZ8	Q9daz8 mus musculu
8	41	57.7	332	11 Q91ZD5	Q91zd5 mus musculu
9	41	57.7	333	11 Q91Z75	Q91z75 mus musculu
10	41	57.7	511	16 Q8Z462	Q8z462 salmoneilla
11	41	57.7	2348	5 Q9V346	Q9v346 drosophila
12	40	56.3	306	4 Q8TAX2	Q8tax2 homo sapien
13	40	56.3	326	4 Q9NX36	Q9nx36 homo sapien
14	40	56.3	338	16 Q84824	Q84824 chlamydia t
15	40	56.3	431	3 Q12124	Q12124 saccharomyc
16	40	56.3	466	4 Q9NUS2	Q9nus2 homo sapien

Q9vii8 drosophila  
Q9xz30 drosophila  
Q8sct3 pseudomonas  
Q97ti8 clostridium  
Q96m1 homo sapien  
Q8xae2 ralatonia s  
Q9ewb1 streptomyce  
Q9ewb4 streptomyce  
P79722 brachydanio  
Q9vbk9 drosophila  
Q960c1 drosophila  
Q9nzw0 homo sapien  
Q9nzw3 homo sapien  
Q38115 bacterioph  
Q96rs9 homo sapien  
Q949el oryza sativ  
Q94937 homo sapien  
Q8u788 agrobacteri  
Q9xvd1 caenorhabdi  
Q9bnx0 unidentified  
Q9bnw0 peripatus s  
Q9bnw7 scolopendra  
Q945c7 hordeum vul  
Q9bme7 aedes aegyp  
Q95p39 aedes aegyp  
Q8t4s0 aedes aegyp  
Q8t4s9 aedes aegyp  
Q9m7a9 oryza sativ  
Q8vwn1 triticum ae

## ALIGNMENTS

### RESULT 1

Q8TCD2 Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Rutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAP CRC64;

Query Match 85.9%; Score 61; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. NO. 0.0091;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNMRRKV 11

Db 39 CFQWQNMRRKV 49

### RESULT 2

Q9UCY5 Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96081613; PubMed=8551695;

RA Sato I.;

RT "Characterization of the 84-kDa protein with ABH activity in human

semenal plasma.";

RL Jpn. J. Legal Med. 49:281-293 (1995).

DR HSSP; P02788; 1BXA.

DR InterPro; IPR001156; Transferrin.

DR Pfam; PF00405; transferrin; 1.

SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDDDB CRC64;

Query Match

Best Local Similarity 80.3%; Score 57; DB 4; Length 38;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FQWQENMRKV 12

Db 21 FQWQENMRKV 31

RESULT 3

ID Q9TR80

AC Q9TR80; PRELIMINARY; PRT; 33 AA.

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE Lactoferrin (Fragment).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID=9940;

RN [1]

RP SEQUENCE.

RX MEDLINE=95127729; PubMed=7827104;

RA Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;

RL Biochim. Biophys. Acta 1243:25-32 (1995).

DR HSSP; O77698; 1CE2.

DR InterPro; IPR001156; Transferrin.

DR Pfam; PF00405; transferrin; 1.

SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match

Best Local Similarity 70.4%; Score 50; DB 6; Length 33;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQENMRKV 11

Db 19 CFQWQENMRKV 29

RESULT 4

ID Q9CGN2

AC Q9CGN2; PRELIMINARY; PRT; 289 AA.

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE Hypothetical 33.3 kDa protein.

GN F28L5.13.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=21016719; PubMed=11130712;

RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,  
 RA Mitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:816-820 (2000).  
 DR EMBL; AC079280; AAC50577.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 289 AA; 33338 MW; 753AA27BED0F840C CRC64;

Query Match 59.2%; Score 42; DB 10; Length 289;

Best Local Similarity 58.3%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWQENMRKV 12

Db 8 CFQWQENMRKV 19

RESULT 5

ID Q90863

AC Q90863; PRELIMINARY; PRT; 81 AA.

DT 01-NOV-1998 (TRENBLrel. 08, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE Glycoprotein gp120 (Fragment).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NJS182;

RA Brandful J.A.M., Ampofo W.K., Janssens W., Adu-Sarkodie Y.,

RA Apeagyei F., Anyomi F., Aidoo S., Barnor J.S., Yamamoto N.,

RA Ishikawa K., Sata T., Kurata T.;

RT "Genetic and phylogenetic analysis of HIV-1 strains from Southern

Ghana.";

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ225659; CAA12541.1; --

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

KW AIDS; Coat protein; Glycoprotein.

FT NON\_TER 1 81

SQ SEQUENCE 81 AA; 9138 MW; 2D43DCD554295572 CRC64;

Query Match

Best Local Similarity 57.7%; Score 41; DB 15; Length 81;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QWQENMRKV 11

Db 69 EWKENLRKV 77

RESULT 6

Q93780

ID Q93780

PRELIMINARY; PRT; 275 AA.

AC Q93780;  
 DT 01-FEB-1997 (TRENBLrel. 02, Created)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE F53H4.4 protein.  
 GN F53H4.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]\_SEQUENCE FROM N.A.  
 RA Dobson R.;  
 RP SEQUENCE FROM N.A.  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81089; CAB03137.1; -.  
 SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC7040DB2 CRC64;  
 Query Match 57.7%; Score 41; DB 5; Length 275;  
 Best Local Similarity 63.6%; Pred. No. 15;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FQWENMKVR 12  
 DB 262 FQWISMKR 272  
 ID Q9DAZ8 PRELIMINARY; PRT; 332 AA.  
 AC Q9DAZ8;  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE 1600000123Rik protein.  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PLACENTA;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach K., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK005389; BAB23995.1; -.  
 DR HSSP; P07711; 1CUL.  
 DR MEOPS; C01.053; -.

DR MGI:1916256; 1600000123Rik.  
 DR InterPro; IPR000668; Peptidase C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF00112; Peptidase C1; 1.  
 DR PRINTS; PD00705; PAPA1N.  
 DR ProDom; PD000158; Peptidase C1; 1.  
 DR ProSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
 SQ SEQUENCE 332 AA; 37298 MW; 0804F1BA5B6538E0 CRC64;  
 Query Match 57.7%; Score 41; DB 11; Length 332;  
 Best Local Similarity 55.6%; Pred. No. 19;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WQENMKVR 12  
 DB 52 WEENMKIK 60  
 ID Q91ZD5 PRELIMINARY; PRT; 332 AA.  
 AC Q91ZD5;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Cathepsin-3 precursor.  
 GN CT53.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PLACENTA;  
 RA Deussing J., Koudio M., Rehman S., Werber I., Schwinde A., Peters C.;  
 RT "Identification and Characterization of a Dense Cluster of Placenta-  
 RT specific Cysteine Peptidases and Related Genes on Mouse Chromosome  
 RT 13."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY034574; AAK58450.1; -.  
 DR MGI:2151929; Cts3.  
 DR InterPro; IPR000668; Peptidase C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF00112; Peptidase C1; 1.  
 DR ProDom; PD000158; Peptidase C1; 1.  
 DR ProSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
 KW SIGNAL.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 114 332 CATHEPSIN-3.  
 SQ SEQUENCE 332 AA; 37326 MW; 4184B90725B41C0D CRC64;  
 Query Match 57.7%; Score 41; DB 11; Length 332;  
 Best Local Similarity 55.6%; Pred. No. 19;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WQENMKVR 12  
 DB 52 WEENMKIK 60  
 ID Q91Z75 PRELIMINARY; PRT; 333 AA.  
 AC Q91Z75;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Cathepsin M.  
 GN CTSM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

```

RN SEQUENCE FROM N.A.
RP STRAIN=129/SEVITACFR; TISSUE=SPLEEN;
RA Rehman S., Peters C., Deussing J.;
RL EMBL; AY057446; AAL15416.1; -
DR InterPro; IPR000668; Peptidase C1.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PROSITE; PS00639; THIOLESTERASE_HIS; UNKNOWN 1.
SQ SEQUENCE 333 AA; 37388 MW; 6DD0BEB91C033110 CRC64;

Query Match 57.7%; Score 41; DB 11; Length 333;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 WOENMKRVR 12
DB 52 WEENMKKIK 60

RESULT 10
ID Q82462 PRELIMINARY; PRT; 511 AA.
AC Q82462;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein STV3070.
GN STV3070.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN SEQUENCE FROM N.A.
RP STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RL Nature 413:848-852(2001).
DR EMBL; AL627276; CAD06049.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DDDI24E10D178B CRC64;

Query Match 57.7%; Score 41; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQENMKRVR 12
DB 350 CFQWQENMKRVR 361

RESULT 11
QYV346 PRELIMINARY; PRT; 2348 AA.
AC QYV346;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG8723 protein.
GN CG11198 OR CG8723.
OS Drosophila melanogaster (Fruit fly).

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfankoch C., Beasley E.M.,
RA Balow R.M., Basu A., Baxendale J., Bayraktarglu L., Bolshakov S.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Heston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasegman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AE003839; AAF59156.1; -
DR HSP; P24182; 1DV1.
DR FlyBase; FBgn0033246; CG111198.
DR InterPro; IPR001882; Biotin attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR000901; Cpsase.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; Cpsase_L_chain; 1.
DR Pfam; PF02786; Cpsase_L_D2; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00866; CFSASE_1; 1.
DR PROSITE; PS00867; CFSASE_2; 1.
KW Biotin.
SQ SEQUENCE 2348 AA; 263722 MW; ED7B6FB9976E1CD2 CRC64;

Query Match 57.7%; Score 41; DB 5; Length 2348;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQENMKRVR 12
DB 563 CFQWQENMKRVR 574

```



```

RESULT 12
Q8TAX2          PRELIMINARY;      PRT;      306 AA.
AC Q8TAX2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to hypothetical protein FLJ11175.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025708; AAH25708.1; -.
KW Hypothetical protein.
SQ SEQUENCE 306 AA; 34551 MW; 43D9AD722566B343 CRC64;

Query Match          56.3%; Score 40; DB 4; Length 306;
Best Local Similarity 55.8%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQENMR 9
   ||||:  :
Db 269 CFQWESTLR 277

RESULT 13
Q9NX36          PRELIMINARY;      PRT;      326 AA.
AC Q9NX36;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ20461 fis, clone KAT06105.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isoigai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000468; BAA91185.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF02226; DnaJ; 1.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS50076; DnaJ2; 1.
SQ SEQUENCE 326 AA; 38324 MW; D6C78F8AABE2243E CRC64;

Query Match          56.3%; Score 40; DB 4; Length 326;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQENMRKV 11
   ||||:  :
Db 297 CFQWQENIRKL 307

RESULT 14
O84824          PRELIMINARY;      PRT;      398 AA.
AC O84824;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

```

```

DE Tyrosine transport protein.
GN TYRP_1 OR CT617.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UM-3/CX; PubMed=9784136;
RX MEDLINE=99000809;
RA Stephens R.S., Kaiman S., Tammul C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
DR EMBL; AE001354; AAC68414.2; -.
DR InterPro; IPR002422; AA/rel_primease2.
DR InterPro; IPR002091; AAA_permease.
DR Pfam; PF03222; TYP_Typ_perm; 1.
DR PRINTS; PR00166; AROAARMEASE.
DR TIGRFAMs; TIGR00837; araap; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 44266 MW; 45B208AE61714A5E CRC64;

Query Match          56.3%; Score 40; DB 16; Length 398;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWQENMRKV 11
   ||||:  :
Db 300 FQWDEKKRV 309

RESULT 15
Q12124          PRELIMINARY;      PRT;      431 AA.
AC Q12124;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chromosome IV reading frame ORF YDL005C.
GN MED2 OR D2930 OR YDL005C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Urrestarazu L.A., Andre B., Vissers S.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Murphy L., Richards C., Gentles S., Harris D.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Barrell B., Rajandream M.A.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=ALPHA S288C;
RA Andre B., Vissers S., Urrestarazu L.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z74053; CAA98561.1; -.
DR EMBL; Z48008; CAA88056.1; -.
DR EMBL; Z48432; CAA88354.1; -.
DR SGD; S0002163; MED2.

```

KW Hypothetical protein.  
SQ SEQUENCE 431 AA; 47717 MW; D3B0992B1E1A4892 CRC64;  
Query Match 56.3%; Score 40; DI 3; Length 431;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 QENMRKVR 12  
| | | | |  
Db 122 QENMRKVR 129

Search completed: February 21, 2003, 07:44:39  
Job time : 21.8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 28.3 seconds  
(without alignments)  
56.502 Million cell updates/sec

Title: US-09-743-107B-86  
Perfect score: 70  
Sequence: 1 CFQWQEMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	12	21 AAY78086	Human lactoferrin
2	65	92.9	12	21 AAY78038	Human lactoferrin
3	65	92.9	12	21 AAY78046	Human lactoferrin
4	65	92.9	12	21 AAY78047	Human lactoferrin
5	65	92.9	13	21 AAY78037	Human lactoferrin
6	65	92.9	13	21 AAY78048	Human lactoferrin
7	65	92.9	13	21 AAY78049	Human lactoferrin
8	65	92.9	14	21 AAY78036	Human lactoferrin
9	65	92.9	14	21 AAY78050	Human lactoferrin
10	65	92.9	14	21 AAY78051	Human lactoferrin

11	65	92.9	15	17 AAR98554	Peptide for anti-u
12	65	92.9	15	21 AAY78035	Human lactoferrin
13	65	92.9	15	21 AAY78062	Human lactoferrin
14	65	92.9	15	21 AAY78063	Human lactoferrin
15	65	92.9	16	21 AAY78031	Human lactoferrin
16	65	92.9	16	21 AAY78064	Human lactoferrin
17	65	92.9	16	21 AAY78065	Human lactoferrin
18	65	92.9	17	21 AAY78034	Human lactoferrin
19	65	92.9	17	21 AAY78066	Human lactoferrin
20	65	92.9	17	21 AAY78067	Human lactoferrin
21	65	92.9	18	15 AAR69352	Human lactoferrin
22	65	92.9	18	17 AAW13397	Advanced glycosyla
23	65	92.9	18	21 AAY78033	Human lactoferrin
24	65	92.9	19	21 AAY68867	Human lactoferrin
25	65	92.9	19	21 AAY78032	Amino acid sequenc
26	65	92.9	20	13 AAR21810	Human lactoferrin
27	65	92.9	20	14 AAR48441	Anti microbial pep
28	65	92.9	20	15 AAR48530	Lactoferrin-relate
29	65	92.9	20	15 AAR48531	Lactoferrin derive
30	65	92.9	20	15 AAR57461	Lactoferrin derive
31	65	92.9	20	15 AAR57462	Lactoferrin derive
32	65	92.9	20	16 AAR84698	Bovine lactoferrin
33	65	92.9	20	16 AAR84699	Bovine lactoferrin
34	65	92.9	20	16 AAR80263	Anti-parasitic lac
35	65	92.9	20	16 AAR80264	Anti-parasitic lac
36	65	92.9	20	17 AAR98553	Peptide for anti-u
37	65	92.9	20	17 AAR91852	Lactoferrin-derive
38	65	92.9	20	17 AAW03045	Lactoferrin-derive
39	65	92.9	20	17 AAR90607	Lactoferrin derive
40	65	92.9	20	17 AAR87621	Lactoferrin-derive
41	65	92.9	20	17 AAR87622	Lactoferrin-derive
42	65	92.9	20	18 AAW26150	Lactoferrin deriva
43	65	92.9	20	18 AAW14036	Anti-parasitic pep
44	65	92.9	20	19 AAW70310	Thrombus formation
45	65	92.9	20	19 AAW53224	Lactoferrin hydrol

## ALIGNMENTS

RESULT 1  
AAY78086  
ID AAY78086 standard; Peptide; 12 AA.  
XX  
AC AAY78086;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:86.

Human; lactoferrin; modification; infection; inflammation; tumour;  
food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
urinary tract infection; colitis; Candida infection; fungicidal;  
bactericidal; preservative.

OS Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

DR

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumors. The peptides can also be used  
CC in food stuffs as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;  
SQ Query Match 100.0%; Score 70; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12  
Db 1 CFQWQREMRKVR 12  
|||||

RESULT 2  
AAY78038  
ID AAY78038 standard; Peptide; 12 AA.

XX AAY78038;  
AC AAY78038;  
DT 25-APR-2000 (first entry)  
DE Human lactoferrin derived peptide SEQ ID NO:38.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.  
PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 12; Page 70; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumors. The peptides can also be used  
CC in food stuffs as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;  
SQ Query Match 92.9%; Score 65; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00018;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12  
Db 1 CFQWQREMRKVR 12  
|||||

RESULT 3  
AAY78046  
ID AAY78046 standard; Peptide; 12 AA.

XX AAY78046;  
AC AAY78046;  
DT 25-APR-2000 (first entry)  
DE Human lactoferrin derived peptide SEQ ID NO:46.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 35; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumors. The peptides can also be used  
CC in food stuffs as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 92.9%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00018;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWOREMRKVR 12  
 ||||| |||||  
 Db 1 CFOWOREMRKVR 12

RESULT 4

AAV78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:47.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 92.9%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00018;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWOREMRKVR 12  
 ||||| |||||  
 Db 1 CFOWOREMRKVR 12

RESULT 5

AAV78037  
 ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:37.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 12; Page 70; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 92.9%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.0002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12  
 ||||| |||||  
 Db 2 CFQWQREMRKVR 13

## RESULT 6

AAAY78048  
 ID AAY78048 standard; Peptide; 13 AA.

XX AAY78048;  
 XX  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 92.9%; Score 65; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.0002;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12

||||| |||||

Db 2 CFQWQREMRKVR 13

## RESULT 7

AAAY78049

ID AAY78049 standard; Peptide; 13 AA.

XX AAY78049;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 92.9%; Score 65; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.0002;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12

||||| |||||

Db 2 CFQWQREMRKVR 13

## RESULT 8

AAAY78036

ID AAY78036 standard; Peptide; 14 AA.

XX AAY78036;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCII-) A+ SCI INVEST AB.  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI  
 XX WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 69; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 92.9%; Score 65; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQREMRKVR 12  
 Db 3 CFQWQREMRKVR 14  
 ||||| |||||  
 RESULT 9  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX  
 AC AAY78050;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:50.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX  
 PN W0200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCII-) A+ SCI INVEST AB.  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI  
 XX WPI; 2000-147388/13.  
 DR  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 15; Page 75; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 92.9%; Score 65; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQREMRKVR 12  
 Db 3 CFQWQREMRKVR 14  
 ||||| |||||  
 RESULT 10  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX  
 AC AAY78051;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:51.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF





Query Match 92.9%; Score 65; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00023;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWOREMRKVR 12  
 DB 4 CFQWOREMRKVR 15

## RESULT 13

AA78062  
 ID AAY78062 standard; Peptide; 15 AA.

XX AC AAY78062;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:62.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

PN W0200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 81; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 92.9%; Score 65; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00023;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWOREMRKVR 12  
 DB 4 CFQWOREMRKVR 15

## RESULT 14

AA78063  
 ID AAY78063 standard; Peptide; 15 AA.

XX AC AAY78063;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:63.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

PN W0200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 81; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 92.9%; Score 65; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00023;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWOREMRKVR 12  
 DB 4 CFQWOREMRKVR 15

## RESULT 15

AA78031  
 ID AAY78031 standard; Peptide; 16 AA.

XX XX  
 AC AAY78031;

```

XX 25-APR-2000 (first entry)
XX
XX Human lactoferrin derived peptide SEQ ID NO:31.
XX
XX Human; lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO200001730-A1.
XX
XX 13-JAN-2000.
XX
XX 06-JUL-1999; 99WO-SE01230.
XX
XX 06-JUL-1998; 98SE-0002441.
PR 17-JUL-1998; 98SE-0002562.
PR 29-DEC-1998; 98SE-0004614.
XX
XX (ASCI-) A+ SCI INVEST AB.
XX
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
PI
XX WPI; 2000-147388/13.
XX
XX New peptides used for treatment and prevention of infections,
PT inflammations and tumors and for use in infant formula food -
XX
XX Claim 11; Page 68; 102pp; English.
XX
XX AAY78001 to AAY78100 represent peptides having sequences based on human
CC lactoferrin. The peptides are taken up in the intestine through
CC binding to specific lactoferrin receptors and are then transported
CC through the circulation. A medicinal product of the peptide or fragment
CC can be used for treating and/or prevention of infections (such as
CC urinary tract infections, colitis, and Candida infection on a mucosal
CC membrane), inflammations and/or tumours. The peptides can also be used
CC in food stuffs such as infant formula food. The peptides are also
CC fungicidal and bactericidal and may also be used as preservatives.
CC Even though native human lactoferrin have been shown to have desired
CC anti-inflammatory anti-infectious and anti-tumoural properties they
CC cannot be used clinically on a broad basis because of high production
CC costs. Therefore, provision of peptides based on lactoferrin would
CC enable them to be used for the same purposes as lactoferrin at lower
CC cost.
XX
XX Sequence 16 AA;
SQ
Query Match 92.9%; Score 65; DB 21; Length 16;
Best Local Similarity 91.7%; Pred. No. 0.00024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWOREMRKVR 12
Db 5 CFQWQNRNRKVR 16

```

Search completed: February 21, 2003, 07:37:15  
 Job time : 28.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:59 ; Search time 8.65 Seconds  
(without alignments)  
40.818 Million cell updates/sec

Title: US-09-743-107B-86  
Perfect score: 70  
Sequence: 1 CFQWQREMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pap:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pap:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	92.9	18	1	US-08-204-487-3
2	65	92.9	18	2	US-08-485-948-8
3	65	92.9	18	2	US-08-628-380-8
4	65	92.9	18	2	US-08-475-055-8
5	65	92.9	20	1	US-07-755-161A-3
6	65	92.9	20	1	US-07-891-174-3
7	65	92.9	20	1	US-08-204-487-1
8	65	92.9	20	1	US-08-256-771-24
9	65	92.9	20	1	US-08-256-771-25
10	65	92.9	20	1	US-08-381-984-24
11	65	92.9	20	1	US-08-381-984-25
12	65	92.9	22	4	US-09-508-734-4
13	65	92.9	24	4	US-09-508-734-6
14	65	92.9	25	1	US-07-755-161A-10
15	65	92.9	25	1	US-07-891-174-10
16	65	92.9	25	1	US-08-204-487-7
17	65	92.9	29	4	US-09-508-734-8
18	65	92.9	36	1	US-07-755-161A-8
19	65	92.9	36	1	US-07-891-174-8
20	65	92.9	36	1	US-08-256-771-30
21	65	92.9	36	1	US-08-381-984-29
22	65	92.9	47	2	US-08-464-182A-6
23	65	92.9	47	2	US-08-406-271-6
24	65	92.9	50	2	US-08-693-274A-7
25	65	92.9	52	4	US-09-017-043A-3
26	65	92.9	53	2	US-08-464-182A-5
27	65	92.9	53	2	US-08-406-271-5

28	65	92.9	54	2	US-08-464-182A-2
29	65	92.9	54	2	US-08-406-271-2
30	65	92.9	694	3	US-08-724-586-2
31	65	92.9	694	4	US-08-421-632-2
32	65	92.9	694	4	US-09-932-190-2
33	65	92.9	705	2	US-08-655-640-2
34	65	92.9	708	2	US-08-655-640-4
35	65	92.9	711	1	US-08-154-019-4
36	65	92.9	711	1	US-08-461-333-4
37	65	92.9	711	3	US-08-464-167-4
38	65	92.9	711	3	US-09-158-313-4
39	65	92.9	711	4	US-08-476-798-4
40	62	88.6	711	1	US-08-145-681-2
41	62	88.6	711	1	US-08-250-308-2
42	62	88.6	711	1	US-08-453-703-2
43	62	88.6	711	2	US-08-456-106-2
44	62	88.6	711	3	US-08-456-108-2
45	62	88.6	711	4	US-09-265-577-2

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/082044487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAKKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37) "  
US-08-204-487-3

Query Match 92.9%; Score 65; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 1 CFQWQRMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855982  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,948  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995  
APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-Cl, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 92.9%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 1 CFQWQRMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VIASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,380  
FILING DATE: April 4, 1996  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-Cl, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 92.9%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 1 CFQWQRMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,948
; FILING DATE:
; APPLICATION NUMBER: 08/488,217
; FILING DATE: JUNE 7, 1995
; APPLICATION NUMBER: 08/418,642
; FILING DATE: APRIL 7, 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 947-1-008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; DESCRIPTION: LF-C1, 8-25
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-475-055-8

```

Query Match 92.9%; Score 65; DB 2; Length 18;  
 Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 CFQWQREMRKVR 12
Db 1 CFQWQREMRKVR 12

```

## RESULT 5

```

US-07-755-161A-3
; Sequence 3, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-755-161A-3

```

Query Match

92.9%; Score 65; DB 1; Length 20;

```
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
Db 2 CFQWQREMRKVR 13

RESULT 6
US-07-891-174-3
; Sequence 3, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL LINE:
; CELL LINE:
; ORGANELL:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site

; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-891-174-3

Query Match 92.9%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
Db 2 CFQWQREMRKVR 13

RESULT 7
US-08-204-487-1
; Sequence 1, Application US/08204487
; Patent No. 5565425
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, NAOKI
; APPLICANT: NAKASHIMA, HIDEKI
; APPLICANT: MOSUCHI, WATARU
; APPLICANT: TANAKA, SHIGEAKI
; APPLICANT: DOSAKO, SHUN'ICHI
; APPLICANT: KAWASAKI, YOSHIHIRO
; APPLICANT: UCHIDA, TOSHIKI
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,487
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FUN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
```

```
/ TELEFAX: (617) 248-7100
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..20
/ OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
/ OTHER INFORMATION: DERIVED FROM HUMAN LACTOFERRIN"
US-08-204-487-1

Query Match 92.9%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 2 CFQWQRMNRKVR 13

RESULT 8
US-08-256-771-24
/ Sequence 24, Application US/08256771
/ Patent No. 5656591
/ GENERAL INFORMATION:
/ APPLICANT: Mamoru TOMITA et al.
/ TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
/ TITLE OF INVENTION: PRODUCTS THEREWITH
/ NUMBER OF SEQUENCES: 32
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Wenderoth, Lind & Ponack
/ STREET: 805 Fifteenth Street, N.W., #700
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/256,771
/ FILING DATE: July 22, 1994
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren M. Cheek, Jr.
/ REGISTRATION NUMBER: 33,367
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-8850
/ TELEFAX:
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 24:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: /note= "Cys residues are linked by
/ OTHER INFORMATION: disulfide bond"
```

```
US-08-256-771-24

Query Match 92.9%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 2 CFQWQRMNRKVR 13

RESULT 9
US-08-256-771-25
/ Sequence 25, Application US/08256771
/ Patent No. 5656591
/ GENERAL INFORMATION:
/ APPLICANT: Mamoru TOMITA et al.
/ TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
/ TITLE OF INVENTION: PRODUCTS THEREWITH
/ NUMBER OF SEQUENCES: 32
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Wenderoth, Lind & Ponack
/ STREET: 805 Fifteenth Street, N.W., #700
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/256,771
/ FILING DATE: July 22, 1994
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren M. Cheek, Jr.
/ REGISTRATION NUMBER: 33,367
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-8850
/ TELEFAX:
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 25:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: /note= "Cys residues are protected to
/ OTHER INFORMATION: prevent disulfide bond"
US-08-256-771-25

Query Match 92.9%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 2 CFQWQRMNRKVR 13

RESULT 10
US-08-381-984-24
```

```

; Sequence 24, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine resi
; OTHER INFORMATION: peptides including th
; OTHER INFORMATION: and 19 are bonded by
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified
; OTHER INFORMATION: peptides including th
; OTHER INFORMATION:
US-08-381-984-24
Query Match          92.9%; Score 65; DB
Best Local Similarity 91.7%; Pred.No. 0.00
Matches 11; Conservative 0; Mismatches

QY      1 CFQWQRNWKVR 12
        ||||| |||||
Db       2 CFQWRNWKVR 13

RESULT 11
US-08-381-984-25
; Sequence 25, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
```

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment there  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"  
US-08-381-984-25  
Query Match 92.9%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQRMKVR 12  
Db 2 CFQWQRMKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: KopatentIn 1.71  
SEQ ID NO 4



LENGTH: 22  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-4

Query Match 92.9%; Score 65; DB 4; Length 22;  
Best Local Similarity 91.7%; Pred. No. 0.00012;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
||||| |||||  
DB 2 CFQWQRMKVR 13

RESULT 13  
US-09-508-734-6  
Sequence 6, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 6  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-508-734-6

Query Match 92.9%; Score 65; DB 4; Length 24;  
Best Local Similarity 91.7%; Pred. No. 0.00013;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
||||| |||||  
DB 3 CFQWQRMKVR 14

RESULT 14  
US-07-755-161A-10  
Sequence 10, Application US/07755161A  
Patent No. 5304633  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/755,161A  
FILING DATE: 19910905  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:

INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:

NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"

FEATURE:  
NAME/KEY: modified site  
LOCATION: 21  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"

PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-10

Query Match 92.9%; Score 65; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 0.00014;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
||||| |||||  
DB 4 CFQWQRMKVR 15

```
RESULT 15
US-07-891-174-10
; Sequence 10, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Porack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21

; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
;
; Query Match 92.9%; Score 65; DB 1; Length 25;
; Best Local Similarity 91.7%; Pred. No. 0.00014;
; Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 CFQWQREMRKVR 12
; Db 4 CFQWQREMRKVR 15
;
; Search completed: February 21, 2003, 07:50:37
; Job time : 8.7 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:51:54 ; Search time 10.5 Seconds  
(without alignments)  
35.508 Million cell updates/sec

Title: US-09-743-107B-86  
Perfect score: 70  
Sequence: 1 CFQWQRMKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	92.9	15	9	US-09-798-869-2
2	65	92.9	25	9	US-09-798-869-20
3	65	92.9	694	9	US-10-023-096-2
4	57	81.4	15	9	US-09-798-869-6
5	54	77.1	15	9	US-09-798-869-3
6	54	77.1	25	9	US-09-798-869-23
7	50	71.4	15	9	US-09-798-869-4
8	50	71.4	25	9	US-09-798-869-22
9	47	67.1	15	9	US-09-798-869-8
10	45	65.7	15	9	US-09-798-869-7
11	42	60.0	15	9	US-09-798-869-29
12	42	60.0	15	9	US-09-798-869-30
13	38	54.3	86	9	US-09-738-626-5715
14	38	54.3	489	9	US-09-888-320-2
15	37	52.9	21	10	US-09-864-761-47985
16	36	51.4	46	10	US-09-864-761-48879
17	36	51.4	846	9	US-10-051-409-4
18	35	50.0	15	9	US-09-798-869-5
19	35	50.0	157	10	US-09-867-550-1340

20	35	50.0	209	10	US-09-904-536-8	Sequence 8, Appli
21	35	50.0	209	10	US-09-904-536-9	Sequence 9, Appli
22	35	50.0	209	10	US-09-904-536-11	Sequence 11, Appl
23	35	50.0	209	10	US-09-904-536-12	Sequence 12, Appl
24	35	50.0	209	10	US-09-904-536-13	Sequence 13, Appl
25	35	50.0	209	10	US-09-904-536-14	Sequence 14, Appl
26	35	50.0	209	10	US-09-904-536-15	Sequence 15, Appl
27	35	50.0	209	10	US-09-904-536-16	Sequence 16, Appl
28	35	50.0	209	10	US-09-904-536-17	Sequence 17, Appl
29	35	50.0	209	10	US-09-904-536-18	Sequence 18, Appl
30	35	50.0	212	10	US-09-904-536-10	Sequence 10, Appl
31	35	50.0	235	9	US-10-095-449-6	Sequence 6, Appli
32	35	50.0	235	10	US-09-448-378-1	Sequence 1, Appli
33	35	50.0	235	10	US-09-983-808-6	Sequence 6, Appli
34	35	50.0	235	10	US-09-904-536-1	Sequence 7, Appli
35	35	50.0	360	9	US-09-829-378-7	Sequence 10, Appl
36	35	50.0	388	10	US-09-989-861-8	Sequence 8, Appli
37	34	48.6	95	10	US-09-764-864-1031	Sequence 1031, Ap
38	34	48.6	255	9	US-09-738-626-5071	Sequence 5071, Ap
39	34	48.6	260	9	US-09-738-626-5416	Sequence 5416, Ap
40	34	48.6	264	10	US-09-788-657-24	Sequence 24, Appl
41	34	48.6	264	10	US-09-788-657-25	Sequence 25, Appl
42	34	48.6	269	9	US-09-895-913A-154	Sequence 154, App
43	34	48.6	356	9	US-09-893-519A-64	Sequence 64, Appl
44	34	48.6	398	9	US-09-895-913A-260	Sequence 260, App
45	34	48.6	1212	9	US-10-219-248-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 92.9%; Score 65; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 6.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
Db 3 CFQWQRMKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

```

; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-798-869-20

Query_Match      92.9%; Score 65; DB 9; Length 25;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 3 CFQWQREMRKVR 14

RESULT 3
US-10-023-096-2
; Sequence 2, Application US/10023096
; Patent No. US20020160941A1
; GENERAL INFORMATION:
; APPLICANT: Kruzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Gollnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: 10505/P58185C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-10-023-096-2

```

```

Query_Match      92.9%; Score 65; DB 9; Length 694;
Best Local Similarity 91.7%; Pred. No. 0.0025;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 22 CFQWQREMRKVR 33

RESULT 4
US-09-798-869-6
; Sequence 6, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens
; OTHER INFORMATION: sequence)
; US-09-798-869-6

Query_Match      81.4%; Score 57; DB 9; Length 15;
Best Local Similarity 83.3%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQREMRKVR 12
   ||||| |||||
Db 3 CFQWQREMRKVR 14

RESULT 5
US-09-798-869-3
; Sequence 3, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: CAPRINE
; US-09-798-869-3

Query_Match      77.1%; Score 54; DB 9; Length 15;
Best Local Similarity 72.7%; Pred. No. 0.0043;

```

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWQREMRKV 11  
|:|||||:  
Db 3 CYQWQRMRLK 13

RESULT 6  
US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 77.1%; Score 54; DB 9; Length 25;  
Best Local Similarity 72.7%; Pred. No. 0.007;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOWQREMRKV 11  
|:|||||:  
Db 3 CYQWQRMRLK 13

RESULT 7  
US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 71.4%; Score 50; DB 9; Length 15;  
Best Local Similarity 72.7%; Pred. No. 0.02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQREMRKV 11  
|:|||||:  
Db 3 CLRQWQEMRKV 13

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 8  
US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 71.4%; Score 50; DB 9; Length 25;  
Best Local Similarity 72.7%; Pred. No. 0.032;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQREMRKV 11  
|:|||||:  
Db 3 CLRQWQEMRKV 13

RESULT 9  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 67.1%; Score 47; DB 9; Length 15;  
Best Local Similarity 72.7%; Pred. No. 0.062;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQREMRKV 11  
|:|||||:  
Db 3 CLRQWQEMRKV 13



RESULT 14  
US-09-888-320-2  
; Sequence 2, Application US/09888320  
; Publication No. US20030013090A1  
; GENERAL INFORMATION:  
; APPLICANT: Barry III, Clifton E.  
; APPLICANT: DeBarber, Andrea E.  
; APPLICANT: Mdulili, Khisimuzi  
; APPLICANT: Bekker, Linda-Gail  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by The Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
; FILE REFERENCE: 015280-413100US  
; CURRENT APPLICATION NUMBER: US/09/888,320  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US 60/214,187  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: wild-type EtaA monooxygenase (Rv3854C, EthA)  
US-09-888-320-2

Query Match 54.3%; Score 38; DB 9; Length 489;  
Best Local Similarity 54.5%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQREMKV 11  
| : |||||  
Db 253 CQWPRRMKX 263

RESULT 15  
US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL096701.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EST\_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06  
US-09-864-761-47985

Query Match 52.9%; Score 37; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 3.8;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQR 6  
| : |||||  
Db 16 CFQWRR 21

Search completed: February 21, 2003, 08:08:09  
Job time : 10.55 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 9.6 seconds  
(without alignments)  
120.168 Million cell updates/sec

Title: US-09-743-107b-86

Perfect score: 70

Sequence: 1 CFQWQREMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	92.9	711	1 TFHUL	lactotransferrin p
2	54	77.1	708	2 JC2323	lactoferrin - goat
3	52	74.3	33	2 S52107	lactoferrin - sheep
4	50	71.4	707	1 A28438	lactoferrin precur
5	46	65.7	4568	2 T08030	dysenin beta heavy
6	40	57.1	275	2 T22597	hypothetical prote
7	40	57.1	584	2 C84325	hypothetical prote
8	40	57.1	932	2 T28820	hypothetical prote
9	39	55.7	57	2 D81949	hypothetical prote
10	39	55.7	376	2 S67085	hypothetical prote
11	39	55.7	1266	2 AC3154	conjugation protei
12	39	55.7	1277	2 G98133	probable traA prot
13	38	54.3	121	2 AH3147	hypothetical prote
14	38	54.3	275	1 S07442	interleukin-2 rece
15	38	54.3	291	1 D86713	transcription regu
16	38	54.3	306	1 A33654	cell cycle arrest
17	38	54.3	339	2 T09217	protein sam2B - sp
18	38	54.3	489	2 T06555	probable monooxyge
19	38	54.3	515	2 T00510	probable cytochrom
20	38	54.3	536	2 T24218	hypothetical prote
21	38	54.3	543	2 T00513	cytochrome P450 ho
22	38	54.3	572	1 DERTMX	malate dehydrogena
23	38	54.3	572	2 S44415	malate dehydrogena
24	38	54.3	638	2 S04540	methylmalonyl-CoA
25	38	54.3	979	2 T08316	probable ATP-depen
26	38	54.3	4464	2 D87755	protein T21E12.4 [
27	37	52.9	206	2 H97451	pyridoxamine 5'-ph
28	37	52.9	206	2 AB2670	pyridoxamine 5'-ph
29	37	52.9	208	2 AG3441	probable pyridoxam

#### ALIGNMENTS

##### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R:Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R:Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A>Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148,'T',150-422,'C',424-711 <REY>

A:Cross-references: EMBL:X53561; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A>Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S2659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A:Experimental source: placenta

A>Note: sequence extracted from NCBI backbone (NCBIP:122202)

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A>Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A>Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <STI>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-28,'X',30-31 <ST2>



R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1997  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:89001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:G186815; PIDN:AAA86665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SWKPVN' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LTF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-71/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status experimental  
 F;157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 534-537, 647-652/Disulfide bonds: #status experimental

Query Match 92.9%; Score 65; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMKVR 12  
 DB 39 CFQWQREMKVR 50

RESULT 2  
 JUC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JUC2323  
 R;Le Provost, F.; Nocard, M.; Guerin, G.; Martin, G.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: JUC2323; MUID:94380047; PMID:8093048  
 A;Accession: JUC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 77.1%; Score 54; DB 2; Length 708;  
 Best Local Similarity 72.7%; Pred. No. 0.17;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMKV 11  
 DB 38 CYQWQREMKL 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samouri, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 74.3%; Score 52; DB 2; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.018;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQREMKV 11  
 DB 19 CYQWQREMKL 29

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory  
 A;Reference number: A92596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 265, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.4%; Score 50; DB 1; Length 707;  
 Best Local Similarity 72.7%; Pred. No. 0.82;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQREMKV 11  
 DB 37 CLRWQREMKV 47

```

C;Accession: C84325
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, I.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-584 <STO>
A;Cross-references: GB:AE004437; NID:g10581192; PIDN:AAG19967.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1732C

Query Match 57.1%; Score 40; DB 2; Length 584;
Best Local Similarity 41.7%; Pred. No. 37;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQREMRKVR 12
Db 445 CFTWRKDMERRR 456

RESULT 8
T28820
hypothetical protein F07C3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28820
R;Favell, A.; Gattung, S.
Submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid F07C3.
A;Reference number: Z20528
A;Accession: T28820
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-932 <PAV>
A;Cross-references: EMBL:U50308; PIDN:AAC48001.1; GSPDB:GN00023; CESP:F07C3.1
A;Experimental source: strain Bristol N2; clone F07C3
C;Genetics:
A;Gene: CESP:F07C3.1
A;Map position: 5
A;Introns: 14/2; 55/3; 86/2; 159/3; 220/2; 268/3; 327/1; 363/3; 402/3; 463/2; 559/3; 599

Query Match 57.1%; Score 40; DB 2; Length 932;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FOWQREMRKVR 12
Db 579 FQWQSRARLVK 589

RESULT 9
D81949
hypothetical protein NMA1014 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: D81949
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jags, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: D81949
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <PAR>
A;Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84283.1; PID:g737971
A;Experimental source: serogroup A, strain Z2491
C;Genetics:

```

---

```

A;Map position: IX
A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;
3334/3; 3686/3; 3882/3; 4240/3
C;Superfamily: dynein heavy chain, ciliary
C;Keywords: nucleotide binding; P-loop
F;1919-1926/Region: nucleotide-binding motif A (P-loop)
F;2202-2209/Region: nucleotide-binding motif A (P-loop)
F;2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 65.7%; Score 46; DB 2; Length 4568;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQREMRKVR 12
Db 1852 CFQWQSLRYIQ 1863

RESULT 6
T22597
hypothetical protein F53H4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T22597
R;Dobson, R.
Submitted to the EMBL Data Library, October 1996
A;Reference number: Z19587
A;Accession: T22597
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-275 <WIL>
A;Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4
A;Experimental source: clone F53H4
C;Genetics:
A;Gene: CESP:F53H4.4
A;Map position: X
A;Introns: 67/1; 153/1
C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 57.1%; Score 40; DB 2; Length 275;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FOWQREMRKVR 12
Db 262 FQWQSRARLVK 272

RESULT 7
C84325
hypothetical protein Vng1732c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

```

A;Gene: NMA1014

Query Match 55.7%; Score 39; DB 2; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 5.7;  
 Matches 8; Conservative 1; Mismatches 0; Gaps 0;  
 QY 1 CFQWQREMRKV 12  
 || :||| |||  
 Db 21 CFFSRREMGKVR 32

# RESULT 10

S67085  
 hypothetical protein YOR193w - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: hypothetical protein O4797  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
 C;Accession: S67085  
 R;Hughes, B.; Pohl, T.M.  
 A;Reference number: S66685  
 A;Accession: S67085  
 A;Molecule type: DNA  
 A;Residues: 1-376 <HUG>  
 A;Cross-references: EMBL:Z75101; NID:gl431575; PID:e252673; PID:gl431578; GSPDB:GN00015;  
 A;Experimental source: strain S28C  
 C;Genetics:  
 A;Gene: MIPS:YOR193w  
 A;Cross-references: SGD:S0005719  
 A;Map position: 15R

Query Match 55.7%; Score 39; DB 2; Length 376;  
 Best Local Similarity 60.0%; Pred. No. 36;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWQREMRKV 11

|| :||| :||  
 Db 181 FRWLEMKKL 190

# RESULT 11

AC3154  
 conjugation protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C;Accession: AC3154  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193  
 A;Accession: AC3154  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1266 <KUR>  
 A;Cross-references: GB:AE008689; PIDN:AAL45649.1; PID:gl7743373; GSPDB:GN00187  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: traA  
 A;Map position: linear chromosome

Query Match 55.7%; Score 39; DB 2; Length 1266;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQREMRKV 11

|| :||| :||  
 Db 1142 QWQREMRKV 1150

# RESULT 12

G98133  
 probable traA protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
 C;Accession: G98133  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
 A;Reference number: A97359; PMID:11743194  
 A;Accession: G98133  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1277 <KUR>  
 A;Cross-references: GB:AE007870; PIDN:AAK88593.1; PID:gl5158306; GSPDB:GN00170  
 C;Genetics:  
 A;Gene: AGR I.66  
 A;Map position: linear chromosome

Query Match 55.7%; Score 39; DB 2; Length 1277;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQREMRKV 11

|| :||| :||  
 Db 1153 QWQREMRKV 1161

# RESULT 13

AH3147  
 hypothetical protein Atu4804 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C;Accession: AH3147  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AH3147

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-121 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAL45598.1; PID:gl7743317; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu4804

A;Map position: linear chromosome

Query Match 54.3%; Score 38; DB 2; Length 121;  
 Best Local Similarity 54.5%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQREMRKV 11

|| :||| :||  
 Db 14 CLAWQRRNRV 24

# RESULT 14

S07442

N;Alternate names: CD25

C;Species: Bos primigenius taurus (cattle)

C;Date: 11-Sep-1998 #sequence\_revision 11-Sep-1998 #text\_change 22-Jun-1999

C;Accession: S07442

R;Weinberg, A.D.; Shaw, J.; Paetkau, V.; Bleackley, R.C.; Magnuson, N.S.; Reeves, R.; M

Immunology 63, 603-610, 1998

A;Title: Cloning of cDNA for the bovine IL-2 receptor (bovine Tac antigen).

A;Reference number: S07442; MUID:88212503; PMID:2835311

A;Accession: S07442  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-275 <WEI>  
A;Cross-references: EMBL:M20818; NID:G163208; PIDN:AAA51414.1; PID:G163209  
C;Complex: The high affinity receptor is a heterotrimer of alpha, beta, and gamma chains  
C;Function:  
A;Description: receptor for interleukin-2  
A;Pathway: interleukin-2 stimulated growth and differentiation of T cells, B cells, NK cells  
C;Superfamily: interleukin-2 receptor alpha chain; complement factor H repeat homology  
C;Keywords: cytokine receptor; duplication; glycoprotein; T-cell proliferation; transmembrane  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-275/Product: interleukin-2 receptor alpha chain #status predicted <MAT>  
F;22-243/Domain: extracellular #status predicted <EXT>  
F;24-77/Domain: complement factor H repeat homology <PH1>  
F;123-184/Domain: complement factor H repeat homology <PH2>  
F;244-264/Domain: transmembrane #status predicted <TM>  
F;265-275/Domain: intracellular #status predicted <INT>  
F;24-64,51-77,123-168,152-184/Disulfide bonds: #status predicted  
F;80,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.3%; Score 38; DB 1; Length 275;  
Best Local Similarity 50.0%; Pred. No. 40;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQREMRYR 12  
| | | | |  
DB 261 CLTWQWKKNR 272

RESULT 15  
DB6713  
Transcription regulator [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: D86713  
R;Solotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: D86713  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-291 <STO>  
A;Cross-references: GB:AE005176; PID:G12723619; PIDN:AAK04806.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: rnaB

Query Match 54.3%; Score 38; DB 2; Length 291;  
Best Local Similarity 85.7%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WQEMRK 10  
| | | | |  
DB 189 WQEMRR 195

Search completed: February 21, 2003, 07:47:58  
Job time : 10.65 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 4.6 Seconds  
(without alignments)  
108.199 Million cell updates/sec

Title: US-09-743-107B-86  
Perfect score: 70  
Sequence: 1 CFQWQREMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	65	92.9	711	1	TRFL_HUMAN
2	54	77.1	708	1	TRFL_CAMDR
3	54	77.1	708	1	TRFL_CAPHI
4	50	71.4	707	1	TRFL_MOUSE
5	46	65.7	458	1	DYBE_CHIRE
6	40	57.1	635	1	TRFL_HORSE
7	38	54.3	275	1	IL2A_BOVIN
8	38	54.3	292	1	NLA_DROME
9	38	54.3	306	1	BUB2_YEAST
10	38	54.3	572	1	MAOX_HUMAN
11	38	54.3	572	1	MAOX_RAT
12	38	54.3	637	1	MUTA_PROFR
13	38	54.3	783	1	YNR2_CABEL
14	38	54.3	4568	1	DYHC_CABEL
15	37	52.9	275	1	IL2A_SHEEP
16	37	52.9	903	1	SYLM_HUMAN
17	37	52.9	1135	1	PHVC_SORBI
18	36	51.4	146	1	RPOB_LIRAF
19	36	51.4	466	1	DCE_IACLA
20	36	51.4	572	1	MAOX_MOUSE
21	36	51.4	765	1	Y008_HUMAN
22	36	51.4	1213	1	T2D2_DROME
23	35	50.0	57	1	YE56_ARCFU
24	35	50.0	85	1	FMRD_SALTY
25	35	50.0	235	1	FL3L_HUMAN
26	35	50.0	288	1	LEP4_PSEPU
27	35	50.0	359	1	TPSB_CABEL
28	35	50.0	388	1	SH4_HUMAN
29	35	50.0	440	1	YEB4_ECOLI
30	35	50.0	457	1	ARLY_AQUAE
31	35	50.0	466	1	DCE_IACLC
32	35	50.0	480	1	YQSI_CABEL
33	35	50.0	496	1	MSSI_SCHPO

34 35 50.0 502 1 C911\_ARATH  
35 35 50.0 663 1 PD15\_HUMAN  
36 35 50.0 704 1 TRFL\_PIG  
37 35 50.0 708 1 TRFL\_BUBBU  
38 35 50.0 817 1 NEB2\_RAT  
39 35 50.0 892 1 RAL6\_SCHPO  
40 35 50.0 962 1 YBX7\_SCHPO  
41 35 50.0 1137 1 PHVC\_ORYSA  
42 35 50.0 1167 1 WCI\_NEUCR  
43 34 48.6 179 1 RK27\_TOBAC  
44 34 48.6 211 1 LOLE\_VIBCH  
45 34 48.6 225 1 PYRE\_CRYNE

## ALIGNMENTS

RESULT 1  
ID TRFL\_HUMAN STANDARD; PRT: 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;  
AC Q96KZ5;  
DT DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;  
DE Lactoferrin B; Lactoferrin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rev M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cho Y.Y.;  
RT Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Conneely O.N.;  
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Liang Q., Jimenez-Flores R., Richardson T.;  
RT "Molecular cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA sequences.";  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RT "cDNA cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RA Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Groj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson N., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253-->methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioïd antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumaiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -!- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Secreted  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53961; CAA37914.1; -  
DR EMBL; U07643; AAB60324.1; -  
DR EMBL; M93150; AAA36159.1; -  
DR EMBL; M83202; AAA59511.1; -  
DR EMBL; M83205; AAA58566.1; -  
DR EMBL; M18642; AAA86665.1; -  
DR EMBL; AF332168; AAG48753.1; -  
DR EMBL; BC015822; AAHL5822.1; -  
DR EMBL; BC015823; AAHL5823.1; -  
DR EMBL; M73700; AAA59479.1; -  
DR EMBL; X52941; CAA37116.1; -  
DR EMBL; U95626; AAB57795.1; -  
DR PIR; S11228; TFHL.  
DR PDB; 1LCF; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LPH; 31-OCT-93.  
DR PDB; 1LFI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1LKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

Query Match 92.9%; Score 65; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00066;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQREMKVR 12  
 DB 39 CFQWQREMKVR 50

RESULT 2

ID	TRFL_CAMDR	STANDARD;	PRT;	708 AA.
AC	Q9TUM0; Q9MZS5;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin).			
GN	LTF.			
OS	Camelus dromedarius (Dromedary) (Arabian camel).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.			
OX	NCBI_TaxID=9838;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Somali; TISSUE=Lactating mammary gland;			
RA	Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;			
RT	"Sequence analysis of camel (Camelus dromedarius) lactoferrin.";			
RL	Int. Dairy J. 9:481-486 (1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;			
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases			
CC	-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).			
CC	-!- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.			
CC	-!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.			

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

ID	TRFL_CAMDR	STANDARD;	PRT;	708 AA.
EMBL;	AJ131674; CAB53387.1; -			
EMBL;	AF165879; AAF82241.1; -			
HSP;	O77811; 1B1X			
InterPro;	IPR001156; Transferrin.			
Protein;	PF00405; transferrin; 2.			
PRINTS;	PR00422; TRANSFERRIN.			
SMART;	SM00094; TR FER; 2.			
PROSITE;	PS00205; TRANSFERRIN 1; 2.			
PROSITE;	PS00206; TRANSFERRIN 2; 2.			
PROSITE;	PS00207; TRANSFERRIN 3; 2.			
Transprot;	Iron transport; Glycoprotein; Metal-binding; Repeat;			
Signal.				
FT	CHAIN	1	19	BY SIMILARITY.
FT	REPEAT	20	708	LACTOTRANSFERRIN.
FT	REPEAT	20	363	1.
FT	REPEAT	364	708	2.
FT	DISULFID	28	64	BY SIMILARITY.
FT	DISULFID	38	55	BY SIMILARITY.
FT	DISULFID	134	217	BY SIMILARITY.
FT	DISULFID	176	192	BY SIMILARITY.
FT	DISULFID	189	200	BY SIMILARITY.
FT	DISULFID	250	264	BY SIMILARITY.
FT	DISULFID	367	399	BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .)  
 FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 R -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 77.1%; Score 54; DB 1; Length 708;  
 Best Local Similarity 75.0%; Pred. No. 0.055;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQREMKVR 12  
 DB 38 CAQWQREMKVR 49

RESULT 3

ID	TRFL_CAMDR	STANDARD;	PRT;	708 AA.
AC	Q29477; Q29479;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin).			
GN	LTF.			
OS	Capra hircus (Goat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Capra.			
OX	NCBI_TaxID=9925;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Lee T., Yu S., Kim S., Lee K., Yu D.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=94380047; PubMed=8093048;			
RA	le Provost F., Nocart M., Guerin G., Martin P.;			
RT	"Characterization of the goat lactoferrin cDNA. Assignment of the relevant locus to bovine U12 syntenic group.";			
RL	Biochem. Biophys. Res. Commun. 203:1324-1332(1994).			
CC	-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE.			
CC	-!- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			





CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U02963; AAA19956.1; -

DR InterPro; IPR004273; Dynein heavy.

DR Pfam; PF03028; Dynein heavy; 1.

DR Motor protein; Microtubules; Dynein; ATP-binding; Flagella;

KW Coiled coil. 277 293 COILED COIL (POTENTIAL).

FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).

FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).

FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).

FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).

FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).

FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).

FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).

FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).

FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).

FT NP\_BIND 1919 1926 ATP (POTENTIAL).

FT NP\_BIND 2202 2209 ATP (POTENTIAL).

FT NP\_BIND 2530 2537 ATP (POTENTIAL).

FT NP\_BIND 2879 2886 ATP (POTENTIAL).

SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;

Query Match 65.7%; Score 46; DB 1; Length 4568;

Best Local Similarity 50.0%; Pred. No. 9.4;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWQREMKRV 12

Db 1852 CFQWQSLRYIQ 1863

|||||:::|:

RESULT 6

TRFL\_HORSE STANDARD; PRT; 695 AA.

ID TRFL\_HORSE

AC 077811;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lactotransferrin precursor (Lactoferrin) (Fragment).

LN LTF.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

NCBI\_TaxID=9796;

RN [1]

RN SEQUENCE FROM N.A.

RP Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;

RA "cDNA sequence of mare lactoferrin.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).

RC TISSUE=Milk;

RX MEDLINE=99296631; PubMed=10366507;

RT Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;

RT "Three-dimensional structure of mare dfferic lactoferrin at 2.6-A

RT resolution.";

RL J. Mol. Biol. 289:303-317(1999).

CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH

CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING

CC OF AN ANION, USUALLY BICARBONATE.

CC -1- SUBUNIT: MONOMER.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.



RESULT 8  
 ID NLIA DROME STANDARD; PRT; 292 AA.  
 AC Q9XZL8; Q9V391;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 GN NLIA OR CG6072.  
 OS Drosophila melanogaster (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McCormick A.V., Goldberg M.L.;  
 RT "Gene required for elongation of Meiosis I spindle in Drosophila females."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Berkley;  
 RC MEDLINE=20136006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,  
 RA Rallev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borchard M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Wang R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 CC -1- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 CC -1- SIMILARITY: BELONGS TO THE DSCRI FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF147700; AAD33987.1; -;  
 DR EMBL; AE003712; AAF55285.1; -;  
 DR FlyBase; FBgn0026629; nla.  
 SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;  
 Query Match 54.3%; Score 38; DB 1; Length 292;  
 Best Local Similarity 54.5%; Pred. No. 14;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 FOWOREMKVR 12  
 ||| |::|  
 Db 150 FOWLRSPRLR 160  
 RESULT 9  
 BUB2\_YEAST STANDARD; PRT; 306 AA.  
 ID BUB2\_YEAST  
 AC P26448;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitotic check point protein BUB2 (Cell cycle arrest protein BUB2).  
 GN BUB2 OR YMR055C OR YMR796.08C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=91330299; PubMed=1651171;  
 RA Hoyt M.A., Totis L., Roberts B.T.;  
 RT "S. cerevisiae genes required for cell cycle arrest in response to  
 RT loss of microtubule function."  
 RL Cell 66:507-517(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C; AB972;  
 RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP CHARACTERIZATION.  
 RC MEDLINE=21385309; PubMed=11493673;  
 RA Lee S.E., Jensen S., Frenz L.M., Johnson A.L., Fesquet D.,  
 RA Johnston L.H.;  
 RT "The Bub2-dependent mitotic pathway in yeast acts every cell cycle and  
 RT regulates cytokinesis."  
 RL J. Cell Sci. 114:2345-2354(2001).  
 CC -1- FUNCTION: Part of a checkpoint which monitors spindle integrity  
 CC and prevents premature exit from mitosis. This cell-cycle arrest  
 CC depends upon inhibition of the G-protein Tem1 by the BFA1/BUB2  
 CC complex.  
 CC -1- SUBUNIT: Interacts with BFA1.  
 CC -1- SUBCELLULAR LOCATION: Spindle poles.  
 CC -1- SIMILARITY: TO S.POMBE CDC16.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M64706; AAA16885.1; -;  
 DR EMBL; Z49703; CAA89765.1; -;  
 DR PIR; A39654; A39654.  
 DR SGD; S0004659; BUB2.  
 DR InterPro; IPR000195; RabGAP\_TBC.  
 DR Pfam; PF00566; TBC; 1.  
 DR SMART; SM00164; TBC; 1.  
 DR Cell cycle; Mitosis.  
 SQ SEQUENCE 306 AA; 35027 MW; ALDDBFB545E81EA3 CRC64;  
 CC

```
Query Match      54.3%; Score 38; DB 1; Length 306;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQREMRK 10
   ||||| : :
DB 108 CFAWTQQR 117

RESULT 10
ID MAOX HUMAN STANDARD; PRT; 572 AA.
AC P48163; O16855; Q9BWX8; Q9UIY4; Q9H1W3;
DT 01-FEB-1996 (Rel. 33, Created)
DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).
GN ME1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90233791; PubMed=2699453;
RA Nikodem V.M., Magnuson M.A., Dozin B., Morioka H.;
RT "Coding nucleotide sequence of rat malic enzyme mRNA and tissue
specific regulation by thyroid hormone.";
EL Endocr. Res. 15:547-564 (1989).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=86111756; PubMed=3753699;
RA Magnuson M.A., Morioka H., Tecce M.F., Nikodem V.M.;
RT "Coding nucleotide sequence of rat liver malic enzyme mRNA.";
RL J. Biol. Chem. 261:1183-1186 (1986).
RN [3]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=8906948; PubMed=3211151;
RA Morioka H., Tenmyson G.E., Nikodem V.M.;
RT "Structural and functional analysis of the rat malic enzyme gene
promoter.";
RL Mol. Cell. Biol. 8:3542-3545 (1988).
RN [4]
RP PARTIAL SEQUENCE.
RX MEDLINE=89296914; PubMed=2740332;
RA Morioka H., Magnuson M.A., Mitsuhashi T., Song M.K.H., Rall J.E.,
RA Nikodem V.M.;
RT "Structural characterization of the rat malic enzyme gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4912-4916 (1989).
CC -!- CATALYTIC ACTIVITY: (S)-malate + NADP(+) = pyruvate + CO(2) +
NADPH.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE MALIC ENZYMES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X77244; CAA54460.1; -
DR EMBL; U43944; AAC50613.1; -
DR EMBL; AL391416; CAC36330.1; -
DR EMBL; AL136970; CAC19505.1; -
DR EMBL; AL049699; CAB52344.1; -
DR Genbank; HGNC:6983; ME1.
DR MIM; 154250; -
DR InterPro; IPR001891; Malic_oxred.
DR Pfam; PF00390; malic; 1.
DR PRINTS; PR00072; MALOXROTASE.
DR PROSITE; PS00331; MALIC_ENZYMES; 1.
KW Oxidoreductase; NADP.
FT NE BIND 301 318 NADP (BY SIMILARITY).
FT CONFLICT 438 438 P -> S (IN REF. 2).
SQ SEQUENCE 572 AA; 64149 MW; EA4C8CB36FC619C CRC64;

Query Match      54.3%; Score 38; DB 1; Length 572;
Best Local Similarity 33.3%; Pred. No. 28;

Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQREMRK 12
   ||||| : : : :
DB 556 CYSWPEEVQKIQ 567

RESULT 11
ID MAOX RAT STANDARD; PRT; 572 AA.
AC P13697;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).
GN ME1 OR MOD1 OR MOD-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90233791; PubMed=2699453;
RA Nikodem V.M., Magnuson M.A., Dozin B., Morioka H.;
RT "Coding nucleotide sequence of rat malic enzyme mRNA and tissue
specific regulation by thyroid hormone.";
EL Endocr. Res. 15:547-564 (1989).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=86111756; PubMed=3753699;
RA Magnuson M.A., Morioka H., Tecce M.F., Nikodem V.M.;
RT "Coding nucleotide sequence of rat liver malic enzyme mRNA.";
RL J. Biol. Chem. 261:1183-1186 (1986).
RN [3]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=8906948; PubMed=3211151;
RA Morioka H., Tenmyson G.E., Nikodem V.M.;
RT "Structural and functional analysis of the rat malic enzyme gene
promoter.";
RL Mol. Cell. Biol. 8:3542-3545 (1988).
RN [4]
RP PARTIAL SEQUENCE.
RX MEDLINE=89296914; PubMed=2740332;
RA Morioka H., Magnuson M.A., Mitsuhashi T., Song M.K.H., Rall J.E.,
RA Nikodem V.M.;
RT "Structural characterization of the rat malic enzyme gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4912-4916 (1989).
CC -!- CATALYTIC ACTIVITY: (S)-malate + NADP(+) = pyruvate + CO(2) +
NADPH.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE MALIC ENZYMES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M26594; AAA41563.1; ALT SEQ.
DR EMBL; M26581; AAA41563.1; JOINED.
DR EMBL; M26582; AAA41563.1; JOINED.
DR EMBL; M26583; AAA41563.1; JOINED.
DR EMBL; M26584; AAA41563.1; JOINED.
DR EMBL; M26585; AAA41563.1; JOINED.
DR EMBL; M26586; AAA41563.1; JOINED.
DR EMBL; M26587; AAA41563.1; JOINED.
DR EMBL; M26588; AAA41563.1; JOINED.
DR EMBL; M26589; AAA41563.1; JOINED.
DR EMBL; M26590; AAA41563.1; JOINED.
```

```

DR EMBL; M26591; AAA41563.1; JOINED.
DR EMBL; M26592; AAA41563.1; JOINED.
DR EMBL; M26593; AAA41563.1; JOINED.
DR PIR; A37228; DERTMX.
DR InterPro; IPR001891; Malic_oxred.
DR Pfam; PF00390; malic; 1.
DR PRINTS; PR00072; MALOXRDXTASE.
DR PROSITE; PS00331; MALIC_ENZYMES; 1.
KW Oxidoreductase; NADP.
FT NP BIND 301 318 NADP (BY SIMILARITY).
SQ SEQUENCE 572 AA; 64002 MW; 7D6EB48F3BA7D95B CRC64;

Query Match 54.3%; Score 38; DB 1; Length 572;
Best Local Similarity 33.3%; Pred. No. 28;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQREMRKV 12
DB 556 CYSWPEEVQKIQ 567

RESULT 12
MUT_A_PROF
ID MUT_A_PROF STANDARD; PRT; 637 AA.
AC F11652;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylmalonyl-CoA mutase small subunit (EC 5.4.99.2) (MCB-beta).
GN MUTA.
OS Propionibacterium freudenreichii shermanii.
OC Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae;
OC Actinomycetales; Propionibacterineae; Propionibacteriaceae;
OC Propionibacterium.
OX NCBI_TaxID=1752;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 3-8; 95-99; 301-305; 472-477
RP AND 522-526.
RC STRAIN=NCIB 9885;
RX MEDLINE=89350816; PubMed=2569861;
RA Marsh E.N., McKie N., Davis N.K., Leadlay P.F.;
RT "Cloning and structural characterization of the genes coding for
RT adenosylcobalamin-dependent methylmalonyl-CoA mutase from
RT Propionibacterium shermanii.";
RL Biochem. J. 260:345-352(1989).
RN [2]
SEQUENCE OF 514-517.
RC STRAIN=NCIB 9885;
RX MEDLINE=89350815; PubMed=2569860;
RA Marsh E.N., Leadlay P.F.;
RT "Methylmalonyl-CoA mutase from Propionibacterium shermanii. Evidence
RT for the presence of two masked cysteine residues.";
RL Biochem. J. 260:339-343(1989).
RN [3]
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RC STRAIN=NCIB 9885;
RX MEDLINE=96398619; PubMed=8805541;
RA Mancía F., Keep N.H., Nakagawa A., Leadlay P.F., McSweeney S.,
RA Rasmussen B., Bosccke P., Diat O., Evans P.R.;
RT "How coenzyme B12 radicals are generated: the crystal structure of
RT methylmalonyl-coenzyme A mutase at 2-A resolution.";
RL Structure 4:339-350(1996).
RN [4]
X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC STRAIN=NCIB 9885;
RX MEDLINE=98322359; PubMed=9655823;
RA Mancía F., Evans P.R.;
RT "Conformational changes on substrate binding to methylmalonyl CoA
RT mutase and new insights into the free radical mechanism.";
RL Structure 6:711-720(1998).
CC -!- FUNCTION: CATALYZES THE ISOMERIZATION OF SUCCINYL-CoA TO
CC METHYLMALONYL-CoA DURING SYNTHESIS OF PROPIONATE FROM
CC TRICARBOXYLIC ACID-CYCLE INTERMEDIATES.

-!- CATALYTIC ACTIVITY: (R)-2-methyl-3-oxopropionyl-CoA = succinyl-
CoA.
-!- COFACTOR: ADENOSYLCOBALAMIN.
-!- PATHWAY: Propionic acid fermentation.
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
-!- SIMILARITY: BELONGS TO THE METHYLMALONYL-CoA MUTASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X14955; CAA33089.1; -.
DR PIR; S04640; S04640.
DR PDB; 1REQ; 27-JAN-97.
DR PDB; 2REQ; 28-JAN-98.
DR PDB; 3REQ; 24-JUN-98.
DR PDB; 4REQ; 13-JAN-99.
DR PDB; 5REQ; 12-AUG-98.
DR PDB; 6REQ; 09-SEP-98.
DR PDB; 7REQ; 16-SEP-98.
DR InterPro; IPR001629; MM_CoA_mutase.
DR InterPro; IPR004608; MmCoA_mut_beta.
DR Pfam; PF01642; MM_CoA_mutase; 1.
DR TIGRFAMs; TIGR00642; mmCoA_mut_beta; 1.
DR PROSITE; PS00544; METMALONYL_COA_MUTASE; 1.
KW Isomerase; Vitamin B12; Cobalt; 3D-structure.
FT INIT MET 0
SQ SEQUENCE 637 AA; 69333 MW; A9D8809A2C4C9150 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 637;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWQREMRKV 11
DB 33 QWREVEKV 41

RESULT 13
YNR2_CABEL
ID YNR2_CABEL STANDARD; PRT; 783 AA.
AC Q21988;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein R13G10.2 in chromosome III.
GN R13G10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
REVIEWS.
RP Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: FAD (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

EMBL; Z35602; CAA84671.2; -.  
 DR WormPep; R13G10.2; CE25088.  
 DR InterPro; IPR002937; Amino oxidase.  
 DR Pfam; PF01593; Amino oxidase; 1.  
 KW Hypothetical protein; Oxidoreductase; Flavoprotein; PAD.  
 FT NP\_BIND 311 366 PAD (ADP PART) (POTENTIAL).  
 SQ SEQUENCE 783 AA; 88799 MW; 80087E96464DC908 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 783;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 CFQWQREMRKV 12  
 DB 540 CIDWGRDRKVK 551

RESULT 14  
 DYHC CAEEL STANDARD; PRT; 4568 AA.  
 AC Q19020;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dynein heavy chain, cytosolic (DYHC).  
 GN DHC-1 OR T21E12.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=96114101; PubMed=8674131;  
 RA Lye R.J., Wilson R.K., Waterston R.H.;  
 RT "Genomic structure of a cytoplasmic dynein heavy chain gene from the  
 nematode *Caenorhabditis elegans*.";  
 RL Cell Motil. Cytoskeleton 32:26-36(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX Du Z., Maggi L.;  
 RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A  
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND  
 CC ORGANELLES ALONG MICROTUBULES.  
 CC -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF  
 CC INTERMEDIATE AND LIGHT CHAINS.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L33260; AAC37251.1; -.  
 DR EMBL; U80440; AAK21472.1; -.  
 DR WormPep; T21E12.4; CE23997  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR004273; Dynein heavy.  
 DR Pfam; PF03028; Dynein heavy; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.

FT DOMAIN 587 652 COILED COIL (POTENTIAL).  
 FT DOMAIN 814 844 COILED COIL (POTENTIAL).  
 FT DOMAIN 1241 1274 COILED COIL (POTENTIAL).  
 FT DOMAIN 1324 1340 COILED COIL (POTENTIAL).  
 FT DOMAIN 1559 1591 COILED COIL (POTENTIAL).  
 FT DOMAIN 1964 1992 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 3132 3229 COILED COIL (POTENTIAL).  
 FT DOMAIN 3339 3432 COILED COIL (POTENTIAL).  
 FT DOMAIN 3707 3739 COILED COIL (POTENTIAL).  
 FT DOMAIN 4359 4386 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1865 1872 ATP (POTENTIAL).  
 FT NP\_BIND 2163 2170 ATP (POTENTIAL).  
 FT NP\_BIND 2537 2544 ATP (POTENTIAL).  
 FT NP\_BIND 2880 2887 ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 521568 MW; 028E52684F381676 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 4568;  
 Best Local Similarity 54.5%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWQREMRKV 11  
 DB 245 CNRWVKRKY 255

RESULT 15  
 IL2A SHEEP STANDARD; PRT; 275 AA.  
 ID IL2A SHEEP  
 AC P26898;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha  
 subunit) (P55) (TAC antigen) (CD25).  
 GN IL2RA.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=T-cell;  
 RA Verhagen A.A.;  
 RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92241682; PubMed=1572550;  
 RA Bujdoso R., Sargan D.R., Williamson M.L., McConnell I.;  
 RT "Cloning of a cDNA encoding the ovine interleukin-2 receptor 55-kDa  
 protein, CD25.";  
 RL Gene 113:283-284(1992).  
 CC -!- FUNCTION: RECEPTOR FOR INTERLEUKIN-2.  
 CC -!- SUBUNIT: NON COVALENT DIMER OF AN ALPHA AND A BETA CHAINS. IL-2R  
 CC EXIST IN 3 DIFFERENT FORMS: A HIGH AFFINITY DIMER, AN INTERMEDIATE  
 CC AFFINITY MONOMER (BETA CHAIN), AND A LOW AFFINITY MONOMER (ALPHA  
 CC CHAIN). THE HIGH AND INTERMEDIATE AFFINITY FORMS ALSO ASSOCIATE  
 CC WITH A GAMMA CHAIN.  
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z11560; CAA77652.1; -.  
 DR EMBL; X60149; CAA42723.1; -.  
 DR EMBL; A19167; CAA01447.1; -.  
 DR PIR; S18910; S18910.

DR PIR; S18899; S18899.  
 DR PIR; JC1113; JC1113.  
 DR HSP; P01589; IILM.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF00084; sushi; 2.  
 DR SMART; SM00032; CCP; 2.  
 KW Transmembrane; Glycoprotein; Receptor; Repeat; Signal; Sushi.  
 FT SIGNAL 1 21  
 FT CHAIN 22 275  
 FT DOMAIN 22 243  
 FT TRANSMEM 244 262  
 FT DOMAIN 263 275  
 FT DOMAIN 23 78  
 FT DOMAIN 122 185  
 FT DISULFID 24 64  
 FT DISULFID 251 77  
 FT DISULFID 123 168  
 FT DISULFID 152 184  
 FT CARBOHYD 80 80  
 FT CONFLICT 166 166  
 SQ SEQUENCE 275 AA; 30904 MW; 1101A2DE5AC5A088 CRC64;

Query Match 52.9%; Score 37; DB 1; Length 275;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFOWQREMRKVR 12  
 Db 261 CLTWQRWKKNR 272

Search completed: February 21, 2003, 07:28:01  
 Job time : 5.6 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:25:55 ; Search time 20.8 seconds  
(without alignments)  
118.873 Million cell updates/sec

Title: US-09-743-107B-86  
Perfect score: 70  
Sequence: 1 CFQWQEMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL.21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	60	85.7	711 4	Q8TCD2
2	56	80.0	38 4	Q9UCY5
3	52	74.3	33 6	Q9TR80
4	40	57.1	148 10	Q9XHP1
5	40	57.1	275 5	Q93780
6	40	57.1	443 16	Q97TQ8
7	40	57.1	584 17	Q9HPA3
8	40	57.1	932 5	Q19153
9	39	55.7	57 16	Q9UV34
10	39	55.7	105 10	Q9XFD5
11	39	55.7	207 10	Q9SML1
12	39	55.7	240 10	Q9SML2
13	39	55.7	329 2	Q9FTY4
14	39	55.7	376 3	Q08580
15	39	55.7	866 10	Q9FHI9
16	39	55.7	1277 16	Q8U6F2

17	39	55.7	1778	5	Q9NE65	Q9ne65 leishmania
18	38	54.3	121	16	Q8U6K3	Q8u6k3 agrobacteri
19	38	54.3	233	4	Q8NVX2	Q8nvx2 homo sapien
20	38	54.3	279	16	Q8XSE2	Q8xse2 raistonia s
21	38	54.3	291	16	Q9CHM1	Q9chm1 lactococcus
22	38	54.3	306	4	Q8TAX2	Q8tax2 homo sapien
23	38	54.3	329	12	Q9QB73	Q9qb73 yaba monkey
24	38	54.3	339	10	Q24366	Q24366 spinacia ol
25	38	54.3	459	4	Q9NZW0	Q9nzw0 homo sapien
26	38	54.3	460	4	Q9NZW3	Q9nzw3 homo sapien
27	38	54.3	466	4	Q9NUS2	Q9nus2 homo sapien
28	38	54.3	477	2	Q9RUL5	Q9rul5 corynebacte
29	38	54.3	489	16	P96223	P96223 mycobacteri
30	38	54.3	489	6	Q9SK28	Q9sk28 macaca fasc
31	38	54.3	512	10	Q22185	Q22185 arabidopsis
32	38	54.3	531	10	Q9LTD4	Q9ltd4 arabidopsis
33	38	54.3	543	10	Q22188	Q22188 arabidopsis
34	38	54.3	632	4	Q94937	Q94937 homo sapien
35	38	54.3	707	6	Q95JR7	Q95jr7 macaca fasc
36	38	54.3	789	2	Q9BY92	Q9ey92 corynebacte
37	38	54.3	864	5	O62582	O62582 encephalito
38	38	54.3	864	5	Q8SRG3	Q8sr93 encephalito
39	38	54.3	864	5	Q8SQ16	Q8sq16 encephalito
40	38	54.3	979	17	O52003	O52003 halobacteri
41	37	52.9	64	10	Q9FXD3	Q9fxd3 arabidopsis
42	37	52.9	91	15	O77855	O77855 human immun
43	37	52.9	91	15	O77856	O77856 human immun
44	37	52.9	108	9	Q8SC55	Q8sc55 stx2 conver
45	37	52.9	205	16	Q986A0	Q986a0 rhizobium 1

#### ALIGNMENTS

#### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
ID Q8TCD2;  
AC Q8TCD2;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Rutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL, BC022347; AAH2347.1; ... 1B9C7EE097C45FAP CRC64;  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAP CRC64;

Query Match 85.7%; Score 60; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.022;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQEMRKV 11  
DB 39 CFQWQEMRKV 49

#### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.  
ID Q9UCY5;  
AC Q9UCY5;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96081613; PubMed=8551695;
RA Sato I.;
RT "Characterization of the 84-kDa protein with ABH activity in human
RL seminal plasma.";
RL Jpn. J. Legal Med. 49:281-293(1995).
DR HSP; P02788; IBAK.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDBE CRC64;

Query Match 80.0%; Score 56; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0059;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FOWQREMRKV 12
Db 21 FOWQREMRKV 31

RESULT 3
Q9TR80 PRELIMINARY; PRT; 33 AA.
AC Q9TR80;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Lactoferrin (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=95127729; PubMed=7827104;
RA Qian Z.Y., Jolles P., Migliore-Samouir D., Fiat A.M.;
RL Biochim. Biophys. Acta 1243:25-32(1995).
DR HSP; O77698; 1CE2.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 33 AA; 3914 MW; D1904CAEL5A73961 CRC64;

Query Match 74.3%; Score 52; DB 6; Length 33;
Best Local Similarity 63.6%; Pred. No. 0.026;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FOWQREMRKV 11
Db 19 FOWQREMRKV 29

RESULT 4
Q9XHP1 PRELIMINARY; PRT; 148 AA.
AC Q9XHP1;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE 2S albumin.
OS Sesamum indicum (Oriental sesame) (Gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-TAINAN 1;
RX MEDLINE=20074970; PubMed=10606554;
RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;

RT "Molecular cloning of 11S globulin and 2S albumin, the two major seed
RT storage proteins in sesame.";
RL J. Agric. Food Chem. 47:4932-4938(1999).
DR EMBL; AF091841; AAD42943.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000617; Napiin.
DR InterPro; IPR001768; Try/amyl_inhbtr.
DR Pfam; PF00234; tryp_alpha_ami1; 1.
DR PRINTS; PR00496; NAFIN.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 57.1%; Score 40; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQREMR 9
Db 54 CFQWQREMR 62

RESULT 5
Q93780 PRELIMINARY; PRT; 275 AA.
ID Q93780;
AC Q93780;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE F53H4.4 protein.
GN F53H4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Dobson R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81089; CAB03137.1; -.
SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 57.1%; Score 40; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FOWQREMRKV 12
Db 262 FOWQREMRKV 272

RESULT 6
Q97TQ8 PRELIMINARY; PRT; 443 AA.
ID Q97TQ8;
AC Q97TQ8;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Xre family DNA-binding domain and TPR repeats containing
DE protein.
GN CAP0040.
OS Clostridium acetobutylicum.
OG Plasmid pSOL1.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatsov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE001438; AK76786.1; -.
DR InterPro; IPR001387; HTH 3.
DR InterPro; IPR001440; TPR.
DR SMART; SM00530; HTH_XRE; 1.
DR SMART; SM00028; TPR; 3.
KW DNA-binding; Plasmid; Complete proteome.
SQ SEQUENCE 443 AA; 52786 MW; 726C5BD19BFB91EA CRC64;

Query Match 57.1%; Score 40; DB 16; Length 443;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQREMRK 10
Db 255 CYQWMKEYQK 264
|:|:|:|:|:|

RESULT 7
Q9HPA3 PRELIMINARY; PRT; 584 AA.
AC Q9HPA3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Vng1732c.
GN Vng1732c.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_TaxID=64091;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005078; AAG19967.1; -.
DR InterPro; IPR001646; Speptide repeat.
DR InterPro; IPR001622; X-channel pore.
DR Pfam; PF00805; Pentapeptide; 2.
KW Complete proteome.
SQ SEQUENCE 584 AA; 65151 MW; 21BF5D5F0486CCC6 CRC64;

Query Match 57.1%; Score 40; DB 17; Length 584;
Best Local Similarity 41.7%; Pred. No. 62;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWQREMRKVR 12
Db 445 CFTWRKDMERK 456
|:|:|:|:|:|

RESULT 8
Q19153 PRELIMINARY; PRT; 932 AA.
ID Q19153
AC Q19153;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 105.1 kDa protein.
GS F07C3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditidae.
CX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Pavello A., Gattung S.;
RT "The sequence of C. elegans cosmid F07C3.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50308; AAG24025.1; -.
DR InterPro; IPR000731; HMGCR/patch_5TM.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 932 AA; 105144 MW; 66680619ADACBFD5 CRC64;

Query Match 57.1%; Score 40; DB 5; Length 932;
Best Local Similarity 63.6%; Pred. No. 99;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FQWQREMRKVR 12
Db 579 FQWQREMRKVR 589
|:|:|:|:|:|

RESULT 9
Q9JUV34 PRELIMINARY; PRT; 57 AA.
AC Q9JUV34;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein NMA1014.
GN NMA1014.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
CX NCBI_TaxID=65699;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jørgensen K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162754; CAB84283.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 57 AA; 6491 MW; A6D8781C29E212C5 CRC64;

```

```

Query Match          55.7%; Score 39; DB 16; Length 57;
Best Local Similarity 66.7%; Pred. No. 8.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWQREMRKVR 12
  |||:|:|
Db 21 CFPSRREMGKVR 32

RESULT 10
Q9XFD5 PRELIMINARY; PRT; 105 AA.
AC Q9XFD5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cytochrome P450 (Fragment)
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN SEQUENCE FROM N.A.
RP TISSUE-PANICLE;
RC Liu J., Yang J.;
RT "Suppression subtractive hybridization (SSH) identified candidate
RT genes that are differentially expressed at rice young panicle.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF140486; AAD29699.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 105 AA; 11912 MW; BOEFCD487E19F9 CRC64;

Query Match          55.7%; Score 39; DB 10; Length 105;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOWQREMRK 10
  |||:|:|
Db 61 CFQWERLGKK 70

RESULT 11
Q9SML1 PRELIMINARY; PRT; 207 AA.
AC Q9SML1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cytochrome P450 monooxygenase (Fragment)
GN CYP81E4.
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX NCBI_TaxID=3827;
RN SEQUENCE FROM N.A.
RP STRAIN=CV. ILC 3279; TISSUE=CELL SUSPENSION CULTURE;
RC PubMed=10773344;
RA Overkamp S., Hein F., Barz W.;
RT "Cloning and characterization of eight cytochrome P450 cDNAs from
RT chickpea (Cicer arietinum L.) cell suspension cultures.";
RL Plant Sci. 155:101-108(2000).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AJ249801; CAB56743.1; -.
DR InterPro; IPR001128; Cytochrome_P450.

```

---

```

DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 207 AA; 23476 MW; DE1D9AAAC2D1BFB CRC64;

Query Match          55.7%; Score 39; DB 10; Length 207;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQRE 7
  |||:|:|
Db 165 CFQWKE 171

RESULT 12
Q9SML2 PRELIMINARY; PRT; 240 AA.
AC Q9SML2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cytochrome P450 monooxygenase (Fragment)
GN CYP81E5.
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX NCBI_TaxID=3827;
RN SEQUENCE FROM N.A.
RP STRAIN=CV. ILC 3279; TISSUE=CELL SUSPENSION CULTURE;
RC PubMed=10773344;
RA Overkamp S., Hein F., Barz W.;
RT "Cloning and characterization of eight cytochrome P450 cDNAs from
RT chickpea (Cicer arietinum L.) cell suspension cultures.";
RL Plant Sci. 155:101-108(2000).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AJ249800; CAB56742.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 240 AA; 27150 MW; 19CB488BE67C0407 CRC64;

Query Match          55.7%; Score 39; DB 10; Length 240;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFOWQRE 7
  |||:|:|
Db 198 CFQWKE 204

RESULT 13
Q9F7Y4 PRELIMINARY; PRT; 329 AA.
AC Q9F7Y4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Rep.
GN REP.
OS Salmonella enteritidis.
OC Plaemid PJ.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Rychlik I., Sebikova A.;  
 RT "Plasmid pJ of *Salmonella enteritidis* encoding hypothetical rep  
 protein.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF266389; AAG17348.1; --  
 DR InterPro; IPR000989; Rep.  
 DR Pfam; PF01446; Rep; 1.  
 KW Plasmid.  
 SQ SEQUENCE 329 AA; 37819 MW; C3D8A9AC0175155D CRC64;

Query Match 55.7%; Score 39; DB 2; Length 329;  
 Best Local Similarity 54.5%; Pred. No. 52;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWQREMRKV 12  
 |:|:|:|:|  
 Db 308 FEWTEVRKYR 318

## RESULT 14

Q08580 PRELIMINARY; PRT; 376 AA.  
 ID Q08580  
 AC Q08580  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE ORF YOR193W.  
 GN YOR193W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hughes B., Pohl T.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z75101; CAA39406.1; --  
 DR SGD; S0005719; YOR193W.  
 SQ SEQUENCE 376 AA; 44131 MW; 976FDBCA43AB4F80 CRC64;

Query Match 55.7%; Score 39; DB 3; Length 376;  
 Best Local Similarity 60.0%; Pred. No. 60;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWQREMRKV 11  
 |:|:|:|:|  
 Db 181 FWLREMKKL 190

## RESULT 15

Q09FH9 PRELIMINARY; PRT; 866 AA.  
 ID Q09FH9  
 AC Q09FH9  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Genomic DNA, chromosome 5, pl clone:MFC19.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopses.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.

RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT Fl and TAC clones.";  
 RL DNA Res. 6:183-195(1999).  
 DR EMBL; AB018113; BAB09175.1; --  
 SQ SEQUENCE 866 AA; 100187 MW; 164FAB9B4976D27D CRC64;

Query Match 55.7%; Score 39; DB 10; Length 866;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQREMRKV 12  
 |:|:|:|:|  
 Db 389 CFNWLKFKLR 400

Search completed: February 21, 2003, 07:44:40  
 Job time : 21.8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-87  
Perfect score: 67  
Sequence: 1 CFQWRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A: Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	98.5	12	21 AAY78072	Human lactoferrin
2	65	97.0	12	21 AAY78038	Human lactoferrin
3	65	97.0	12	21 AAY78046	Human lactoferrin
4	65	97.0	12	21 AAY78047	Human lactoferrin
5	65	97.0	12	21 AAY78084	Human lactoferrin
6	65	97.0	12	21 AAY78087	Human lactoferrin
7	65	97.0	12	21 AAY78088	Human lactoferrin
8	65	97.0	12	21 AAY78091	Human lactoferrin
9	65	97.0	12	21 AAY78092	Human lactoferrin
10	65	97.0	13	21 AAY78037	Human lactoferrin

11	65	97.0	13	21 AAY78048	Human lactoferrin
12	65	97.0	13	21 AAY78049	Human lactoferrin
13	65	97.0	14	21 AAY78036	Human lactoferrin
14	65	97.0	14	21 AAY78050	Human lactoferrin
15	65	97.0	14	21 AAY78051	Human lactoferrin
16	65	97.0	15	17 AAR98554	Peptide for anti-u
17	65	97.0	15	21 AAY78035	Human lactoferrin
18	65	97.0	15	21 AAY78082	Human lactoferrin
19	65	97.0	15	21 AAY78063	Human lactoferrin
20	65	97.0	16	21 AAY78031	Human lactoferrin
21	65	97.0	16	21 AAY78064	Human lactoferrin
22	65	97.0	16	21 AAY78065	Human lactoferrin
23	65	97.0	17	21 AAY78034	Human lactoferrin
24	65	97.0	17	21 AAY78086	Human lactoferrin
25	65	97.0	17	21 AAY78067	Human lactoferrin
26	65	97.0	18	15 AAR69352	Human lactoferrin
27	65	97.0	18	17 AAW13397	Advanced glycosyla
28	65	97.0	18	21 AAY78033	Human lactoferrin
29	65	97.0	19	21 AAY68667	Amino acid sequenc
30	65	97.0	19	21 AAY78032	Human lactoferrin
31	65	97.0	20	13 AAR21810	Anti microbial pep
32	65	97.0	20	14 AAR44841	Lactoferrin-relate
33	65	97.0	20	15 AAR48530	Lactoferrin-derive
34	65	97.0	20	15 AAR48531	Lactoferrin-derive
35	65	97.0	20	15 AAR57461	Lactoferrin-derive
36	65	97.0	20	15 AAR57462	Lactoferrin-derive
37	65	97.0	20	16 AAR84698	Bovine lactoferrin
38	65	97.0	20	16 AAR84699	Bovine lactoferrin
39	65	97.0	20	16 AAR80263	Anti-parasitic lac
40	65	97.0	20	16 AAR80264	Anti-parasitic lac
41	65	97.0	20	17 AAR98553	Peptide for anti-u
42	65	97.0	20	17 AAR91852	Lactoferrin-derive
43	65	97.0	20	17 AAW03045	Lactoferrin-derive
44	65	97.0	20	17 AAR90607	Lactoferrin-derive
45	65	97.0	20	17 AAR87621	Lactoferrin-derive

## ALIGNMENTS

RESULT 1

AAV78072

ID AAY78072 standard; Peptide; 12 AA.

XX AAY78072;

XX AAY78072;

DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:72.

XX Human, lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.

XX Homo sapiens.

XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

DR

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
PS Claim 22; Page 35; 102pp; English.  
XX  
CC AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
SQ Sequence 12 AA;  
Query Match 98.5%; Score 66; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 3.1e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWXRNMKVR 12  
DB 1 CFQWXRNMKVR 12  
|||||  
RESULT 2  
AAY78038  
ID AAY78038 standard; Peptide; 12 AA.  
XX  
AC AAY78038;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:38.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
DR WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
PS Claim 12; Page 70; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
SQ Sequence 12 AA;  
Query Match 97.0%; Score 65; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWXRNMKVR 12  
DB 1 CFQWXRNMKVR 12  
|||||  
RESULT 3  
AAY78046  
ID AAY78046 standard; Peptide; 12 AA.  
XX  
AC AAY78046;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:46.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
DR WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
PS Claim 15; Page 35; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWQRNRKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:47.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.  
 XX 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 XX 17-JUL-1998; 98SE-0002562.  
 XX 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWQRNRKVR 12

RESULT 5  
 AAY78084  
 ID AAY78084 standard; Peptide; 12 AA.

XX AC AAY78084;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:84.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.  
 XX 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 XX 17-JUL-1998; 98SE-0002562.  
 XX 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX PS Claim 22; Page 36; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWXRNRKVR 12

## RESULT 6

AAV78087  
 ID AAV78087 standard; Peptide; 12 AA.

AC AAV78087;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:87.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 22; Page 37; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 1 CFQWXRNRKVR 12

## RESULT 7

AAV78088

ID AAV78088 standard; Peptide; 12 AA.

XX AAV78088;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:88.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO2000001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 22; Page 37; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 1 CFQWXRNRKVR 12

## RESULT 8

AAV78091

ID AAV78091 standard; Peptide; 12 AA.

XX AAV78091;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:91.



XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX 13-JAN-2000.  
 PD  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCII-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX  
 DR New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 PT Claim 22; Page 38; 102pp; English.  
 PS  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKVR 12  
 DB ||||| |||||  
 1 CFQWXRNRKVR 12  
 RESULT 9  
 AAY78092  
 ID AAY78092 standard; Peptide; 12 AA.  
 XX  
 AC AAY78092;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:92.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.  
 PD  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCII-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX  
 DR New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 PT Claim 22; Page 38; 102pp; English.  
 PS  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKVR 12  
 DB ||||| |||||  
 1 CFQWXRNRKVR 12  
 RESULT 10  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.  
 XX  
 AC AAY78037;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:37.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX 13-JAN-2000.  
 PD  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.

XX (ASCII-) A+ SCI INVEST AB.  
 PA

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI

XX WPI; 2000-147388/13.  
 DR

XX New peptides used for treatment and prevention of infections,  
 CC inflammations and tumors and for use in infant formula food -

XX Claim 12; Page 70; 102pp; English.  
 PS

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 5e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB |||||  
 2 CFQWRNRKVR 13

RESULT 11

ID AAY78048  
 ID AAY78048 standard; Peptide; 13 AA.

XX AAY78048;  
 AC

XX 25-APR-2000 (first entry)  
 DT

XX Human lactoferrin derived peptide SEQ ID NO:48.  
 DE

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 KW

OS Homo sapiens.  
 OS Synthetic.  
 OS

XX WO200001730-A1.  
 PN

XX 13-JAN-2000.  
 PD

XX 06-JUL-1999; 99WO-SE01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR

XX 17-JUL-1998; 98SE-0002562.  
 PR

XX 29-DEC-1998; 98SE-0004614.  
 PR

XX (ASCII-) A+ SCI INVEST AB.  
 PA

XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.  
 PS

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 5e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB |||||  
 2 CFQWRNRKVR 13

RESULT 12

AAY78049

ID AAY78049 standard; Peptide; 13 AA.

XX AAY78049;  
 AC

XX 25-APR-2000 (first entry)  
 DT

XX Human lactoferrin derived peptide SEQ ID NO:49.  
 DE

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 KW

OS Homo sapiens.  
 OS Synthetic.  
 OS

XX WO200001730-A1.  
 PN

XX 13-JAN-2000.  
 PD

XX 06-JUL-1999; 99WO-SE01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR

XX 17-JUL-1998; 98SE-0002562.  
 PR

XX 29-DEC-1998; 98SE-0004614.  
 PR

XX (ASCII-) A+ SCI INVEST AB.  
 PA

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI

XX WPI; 2000-147388/13.  
 DR

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX

PS Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX

SQ Sequence 13 AA;

Query Match 97.0%; Score 65; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 5e-05; 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 2 CFQWQRNRKVR 13

|||||

RESULT 13

AAY78036

ID AAY78036 standard; Peptide; 14 AA.

XX

AC AAY78036;

XX

DT 25-APR-2000 (first entry)

XX

DE Human lactoferrin derived peptide SEQ ID NO:36.

XX

KW Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200001730-A1.

XX

PD 13-JAN-2000.

XX

PF 06-JUL-1999; 99WO-SE01230.

XX

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX

PA (ASCI-) A+ SCI INVEST AB.

XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX

WPI; 2000-147388/13.

XX

PT New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX

Claim 12; Page 69; 102pp; English.

XX

PS AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX

SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 5.4e-05; 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 3 CFQWQRNRKVR 14

|||||

RESULT 14

AAY78050

ID AAY78050 standard; Peptide; 14 AA.

XX

AC AAY78050;

XX

DT 25-APR-2000 (first entry)

XX

DE Human lactoferrin derived peptide SEQ ID NO:50.

XX

KW Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200001730-A1.

XX

PD 13-JAN-2000.

XX

PF 06-JUL-1999; 99WO-SE01230.

XX

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX

PA (ASCI-) A+ SCI INVEST AB.

XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX

WPI; 2000-147388/13.

XX

PT New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX

Claim 15; Page 75; 102pp; English.

XX

PS AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 5.4e-05;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12

Db 3 CFQWQRNMKVR 14

# RESULT 15

AAV78051

ID AAV78051 standard; Peptide; 14 AA.

XX AC AAV78051;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:51.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food

XX PS Claim 18; Page 75; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12

Db 3 CFQWQRNMKVR 14

Search completed: February 21, 2003, 07:56:42

Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-87

Perfect score: 67

Sequence: 1 CFQWXRMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	65	97.0	18	1	US-08-204-487-3
2	65	97.0	18	2	US-08-485-948-8
3	65	97.0	18	2	US-08-628-380-8
4	65	97.0	18	2	US-08-475-055-8
5	65	97.0	20	1	US-07-755-161A-3
6	65	97.0	20	1	US-07-891-174-3
7	65	97.0	20	1	US-08-204-487-1
8	65	97.0	20	1	US-08-256-771-24
9	65	97.0	20	1	US-08-256-771-25
10	65	97.0	20	1	US-08-381-984-24
11	65	97.0	20	1	US-08-381-984-25
12	65	97.0	22	4	US-09-508-734-4
13	65	97.0	24	4	US-09-508-734-6
14	65	97.0	25	1	US-07-755-161A-10
15	65	97.0	25	1	US-07-891-174-10
16	65	97.0	25	1	US-08-406-487-7
17	65	97.0	29	4	US-09-508-734-8
18	65	97.0	36	1	US-07-755-161A-8
19	65	97.0	36	1	US-07-891-174-8
20	65	97.0	36	1	US-08-256-771-30
21	65	97.0	36	1	US-08-381-984-29
22	65	97.0	47	2	US-08-464-182A-6
23	65	97.0	47	2	US-08-406-271-6
24	65	97.0	50	2	US-08-693-274A-7
25	65	97.0	52	4	US-09-017-043A-3
26	65	97.0	53	2	US-08-464-182A-5
27	65	97.0	53	2	US-08-406-271-5

28	65	97.0	54	2	US-08-464-182A-2	Sequence 2, Appli
29	65	97.0	54	2	US-08-406-271-2	Sequence 2, Appli
30	65	97.0	694	3	US-08-724-586-2	Sequence 2, Appli
31	65	97.0	694	4	US-09-421-632-2	Sequence 2, Appli
32	65	97.0	694	4	US-09-932-190-2	Sequence 2, Appli
33	65	97.0	705	2	US-08-655-640-2	Sequence 2, Appli
34	65	97.0	708	2	US-08-655-640-4	Sequence 4, Appli
35	65	97.0	711	1	US-08-154-019-4	Sequence 4, Appli
36	65	97.0	711	1	US-08-461-333-4	Sequence 4, Appli
37	65	97.0	711	3	US-08-464-167-4	Sequence 4, Appli
38	65	97.0	711	3	US-09-158-313-4	Sequence 4, Appli
39	65	97.0	711	4	US-08-476-798-4	Sequence 4, Appli
40	62	92.5	711	1	US-08-145-681-2	Sequence 2, Appli
41	62	92.5	711	1	US-08-250-308-2	Sequence 2, Appli
42	62	92.5	711	1	US-08-453-703-2	Sequence 2, Appli
43	62	92.5	711	2	US-08-456-106-2	Sequence 2, Appli
44	62	92.5	711	3	US-08-456-108-2	Sequence 2, Appli
45	62	92.5	711	4	US-09-265-577-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 97.0%; Score 65; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 585882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LP-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 97.0%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LP-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 97.0%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;;  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LP-CL, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
US-08-475-055-8

Query Match 97.0%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWQRNRKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-3

Query Match 97.0%; Score 65; DB 1; Length 20;

Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
DB 2 CFQWQRNMKVR 13

## RESULT 6

US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: MAMOTO TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:

; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

; LOCATION: 2  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 19"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 19  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 2"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
DB 2 CFQWQRNMKVR 13

## RESULT 7

US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAKOI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000



```

; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
; OTHER INFORMATION: DERIVED FROM HUMAN LACTOFERRIN"
US-08-204-487-1

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"
US-08-204-487-1

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"
US-08-204-487-1
```

```

US-08-256-771-24

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 9
US-08-256-771-25
; Sequence 25, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are protected to
; OTHER INFORMATION: prevent disulfide bond"
US-08-256-771-25

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 10
US-08-381-984-24
```

; Sequence 24, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-24

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
Db 2 CFQWQRNMKVR 13

RESULT 11  
US-08-381-984-25  
; Sequence 25, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack

; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-25

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
Db 2 CFQWQRNMKVR 13

RESULT 12  
US-09-508-734-4  
; Sequence 4, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Samyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; TITLE OF INVENTION: useful microorganism thereof  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KR99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KR1998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 4

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match          97.0%; Score 65; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 4.4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||||
Db 2 CFQWRNMRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: FA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          97.0%; Score 65; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 4.8e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||||
Db 3 CFQWRNMRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match          97.0%; Score 65; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 5e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
   |||||
Db 4 CFQWRNMRKVR 15
```

## RESULT 15

US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

## ; IDENTIFICATION METHOD:

; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10

Query Match 97.0%; Score 65; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred No. 5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNMKKVR 12  
||| |||||  
Db 4 CFQWRNMKKVR 15

Search completed: February 21, 2003, 08:04:25  
Job time : 8.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107B-87  
Perfect score: 67  
Sequence: 1 CFQWXRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08 NEW PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05 NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06 PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07 NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07 PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08 PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09 NEW PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09 PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10 NEW PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10 PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60 NEW PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	97.0	15	9	US-09-798-869-2
2	65	97.0	25	9	US-09-798-869-20
3	65	97.0	694	9	US-10-023-096-2
4	57	85.1	15	9	US-09-798-869-6
5	48	71.6	15	9	US-09-798-869-3
6	48	71.6	25	9	US-09-798-869-23
7	40	59.7	15	9	US-09-798-869-7
8	39	58.2	25	9	US-09-798-869-4
9	39	58.2	25	9	US-09-798-869-22
10	37	55.2	489	9	US-09-888-320-2
11	36	53.7	15	9	US-09-798-869-8
12	36	53.7	15	9	US-09-798-869-29
13	35	52.2	15	9	US-09-798-869-30
14	35	52.2	21	10	US-09-864-761-47985
15	34	50.7	86	9	US-09-738-626-5715
16	34	50.7	95	10	US-09-764-864-1031
17	34	50.7	338	9	US-09-978-295A-119
18	34	50.7	338	9	US-09-978-697-119
19	34	50.7	338	9	US-09-978-192A-119

20	34	50.7	338	9	US-09-999-832A-119	Sequence 119, Appl
21	34	50.7	338	9	US-09-978-189-119	Sequence 119, Appl
22	34	50.7	553	10	US-09-796-753-14	Sequence 14, Appl
23	34	50.7	553	10	US-09-981-649A-6	Sequence 6, Appli
24	34	50.7	553	10	US-09-981-649A-24	Sequence 24, Appl
25	34	50.7	554	10	US-09-981-649A-30	Sequence 30, Appl
26	34	50.7	554	10	US-09-981-649A-32	Sequence 32, Appl
27	34	50.7	556	10	US-09-795-691-2	Sequence 2, Appli
28	34	50.7	559	10	US-09-981-649A-28	Sequence 28, Appl
29	34	50.7	846	9	US-10-051-409-4	Sequence 4, Appli
30	34	50.7	1212	9	US-10-219-248-3	Sequence 3, Appli
31	34	50.7	1212	9	US-10-219-247-3	Sequence 3, Appli
32	34	50.7	1212	10	US-09-855-722-3	Sequence 3, Appli
33	34	50.7	1238	9	US-10-219-248-5	Sequence 5, Appli
34	34	50.7	1238	9	US-10-219-247-5	Sequence 5, Appli
35	34	50.7	1238	10	US-09-855-722-5	Sequence 5, Appli
36	34	50.7	1238	10	US-09-944-849-4	Sequence 4, Appli
37	33	49.3	40	10	US-09-864-761-46393	Sequence 46393, A
38	33	49.3	62	10	US-09-815-242-12129	Sequence 12129, A
39	33	49.3	62	10	US-09-815-242-13026	Sequence 13026, A
40	33	49.3	333	9	US-09-796-753-26	Sequence 26, Appl
41	33	49.3	1258	10	US-09-867-852-107	Sequence 107, Appl
42	33	49.3	26926	9	US-09-759-508B-2	Sequence 2, Appli
43	32	47.8	28	10	US-09-864-761-36085	Sequence 36085, A
44	32	47.8	31	9	US-09-956-206A-13	Sequence 13, Appl
45	32	47.8	344	9	US-09-735-056-33	Sequence 33, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 97.0%; Score 65; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 1.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 3 CFQWXRNRKVR 14

##### RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 97.0%; Score 65; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 2.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
DB 3 CFQWQNRKVR 14

RESULT 3  
US-10-023-096-2

; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; TITLE OF INVENTION: Lactoferrin  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 97.0%; Score 65; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.00066;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
DB 22 CFQWQNRKVR 33

RESULT 4

US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 85.1%; Score 57; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00044;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
DB 3 CFQWQNRKVR 14

RESULT 5

US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 71.6%; Score 48; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.016;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKV 11  
|:|:|:|:|:  
Db 3 CYQWQRNRKL 13

RESULT 6  
US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 71.6%; Score 48; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.026;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKV 11  
|:|:|:|:|:  
Db 3 CYQWQRNRKL 13

RESULT 7  
US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 59.7%; Score 40; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.4;  
Matches 6; Conservative 2; Mismatches 2; Indels 3; Gaps 0;

Qy 1 CFQWXRNRKV 11  
|:|:|:|:|:  
Db 3 CYQWQRNRKL 13

RESULT 8  
US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 58.2%; Score 39; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.6;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKV 11  
|:|:|:|:|:  
Db 3 CLRQWNRKV 13

RESULT 9  
US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 58.2%; Score 39; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 0.97;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKV 11  
|:|:|:|:|:  
Db 3 CLRQWNRKV 13





```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47985
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096701.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EST_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06
US-09-864-761-47985

```

```

Query Match      52.2%; Score 35; DB 10; Length 21;
Best Local Similarity 83.3%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CFQWXR 6
    |||||
Db 16 CFQWRR 21

```

```

RESULT 15
US-09-738-626-5715
; Sequence 5715, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:

```

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5715
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5715

Query Match      50.7%; Score 34; DB 9; Length 86;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWXRNMKVR 12
    |::|::|::|
Db 73 FEYRQLRKIR 83

Search completed: February 21, 2003, 08:11:55
Job time : 7.88372 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107B-87

Perfect score: 67

Sequence: 1 CFQWXRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	97.0	711	1	TFHUL
2	48	71.6	708	2	lactotransferrin p
3	45	67.2	33	3	lactoferrin - goat
4	41	61.2	511	2	lactoferrin - shee
5	39	58.2	275	2	hypothetical prote
6	39	58.2	707	1	hypothetical prote
7	38	56.7	531	2	lactoferrin precu
8	38	56.7	536	2	En/Spm-like transp
9	38	56.7	4568	2	hypothetical prote
10	37	55.2	274	2	dynactin beta heavy
11	37	55.2	489	2	apolipoprotein B-1
12	37	55.2	501	2	probable monooxyge
13	37	55.2	584	2	hypothetical sh3-c
14	36	53.7	124	2	hypothetical prote
15	36	53.7	298	2	F1511.22 (imported
16	36	53.7	298	2	hypothetical prote
17	36	53.7	361	2	probable proteinase
18	36	53.7	365	2	MHC class II histoc
19	36	53.7	428	2	MAP1 protein - myx
20	36	53.7	742	2	histidyl-L-tryptoph
21	36	53.7	749	2	hypothetical prote
22	36	53.7	2700	2	outer capsid prote
23	35	52.2	114	2	protein F21H11.2 [
24	35	52.2	205	2	carcinoembryonic a
25	35	52.2	206	2	26S proteasome SU
26	35	52.2	206	2	pyridoxamine 5'-ph
27	35	52.2	208	2	pyridoxamine 5'-ph
28	35	52.2	376	2	probable pyridoxam
29	35	52.2	649	2	hypothetical prote
30	35	52.2	932	2	hypothetical prote
31	35	52.2	966	1	RNA la protein - b
32	35	52.2	1135	2	phytochrome C - so
33	35	52.2	1174	2	pyruvate carboxyla
34	35	52.2	1174	2	pyruvate carboxyla
35	35	52.2	1213	2	probable helicase
36	35	52.2	1432	2	trichohyalin like
37	35	52.2	1804	2	resistance protein
38	35	52.2	6842	2	protein UNC-89 - C
39	34	50.7	214	2	vif protein - siml
40	34	50.7	224	2	hypothetical prote
41	34	50.7	249	2	arginyltransferase
42	34	50.7	255	2	signal peptidase I
43	34	50.7	289	2	33-3K hypothetical
44	34	50.7	323	2	conserved hypothet
45	34	50.7	335	2	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74

R/Cho, Y.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>

A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>

A/Cross-references: EMBL:X53961; NID:G34415; PID:G34416

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <TEN>

A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A/Experimental source: placenta

A/Note: sequence extracted from NCBI backbone (NCBIP:122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POW>

A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 20-31 <ST1>

A/Accession: S20841

A/Molecule type: protein

A/Residues: 20-28, 'X', 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A' 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:gl86815; PIDN:AAA86665.1; PID:g386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.Y.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SWKPVN' <PAN>  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1984  
 A>Title: Human lactoferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-400  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L. Eur. J. Biochem. 241, 303-308, 1996  
 A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity chromatography  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LTF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status experimental  
 F;155, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #status experimental

Query Match 97.0%; Score 65; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0005;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB 39 CFQWQRNRKVR 50  
 |||||  
 |||||

RESULT 2  
 JC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JC2323  
 R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P. Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: JC2323; MUID:94380047; PMID:8093048  
 A;Accession: JC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.6%; Score 48; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.65;  
 Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 CFQWXRNRKVR 11  
 DB 38 CYQWQRNRKVL 48  
 |||||  
 |||||

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M. Biochim. Biophys. Acta 1243, 25-32, 1995  
 A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 67.2%; Score 45; DB 2; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.12;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 11  
 DB 19 CYQWQRNRKVL 29  
 |||||  
 |||||

RESULT 4  
 AB0858  
 hypothetical protein STV3070 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C;Species: Salmonella enterica subsp. enterica serovar Typhi  
 A;Note: this species has also been called Salmonella typhi  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 C;Accession: AB0858  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.H.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A;Reference number: AB0502; PMID:11677608  
 A;Accession: AB0858  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-511 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:g16504016; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STV3070

Query Match 61.2%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 9.1;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB 350 CFAMDNRKVR 361  
 |||||  
 |||||

RESULT 5  
 T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C;Accession: T22597

R; Dobson, R.  
Submitted to the EMBL Data Library, October 1996  
A; Reference number: Z19587  
A; Accession: T22597  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-275 <WIL>  
A; Cross-references: EMBL:Z81089; PIDN: CAB03137.1; GSPDB: GNO0028; CESP: F53H4.4  
A; Experimental source: clone F53H4  
C; Genetics:  
A; Gene: CESP: F53H4.4  
A; Map position: X  
A; Introns: 67/1; 153/1  
C; Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 58.2%; Score 39; DB 2; Length 275;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWXRNMKVR 12  
| | | | |  
Db 262 FQWKISMKTR 272

RESULT 6  
A28438  
lactoferrin precursor - mouse  
N; Alternate names: lactotransferrin  
C; Species: Mus musculus (house mouse)  
C; Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C; Accession: A28438; A41205  
R; Pentecost, B.T.; Teng, C.T.  
J. Biol. Chem. 262, 10134-10139, 1987  
A; Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory granules  
A; Reference number: A92596; MUID: 87280033; PMID: 3611056  
A; Accession: A28438  
A; Molecule type: mRNA  
A; Residues: 3-707 <PEN>  
A; Cross-references: EMBL: J03298  
R; Liu, Y.; Teng, C.T.  
J. Biol. Chem. 266, 21880-21885, 1991  
A; Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
A; Reference number: A41205; MUID: 92042099; PMID: 1939212  
A; Accession: A41205  
A; Molecule type: DNA  
A; Residues: 1-15 <LIU>  
A; Cross-references: GB: M74778  
C; Superfamily: transferrin; transferrin repeat homology  
C; Keywords: duplication; glycoprotein  
F; 1-19/Domain: signal sequence #status predicted <SIG>  
F; 20-707/Product: lactotransferrin #status predicted <MAT>  
F; 358-695/Domain: transferrin repeat homology <TRH2>  
F; 494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.2%; Score 39; DB 1; Length 707;  
Best Local Similarity 54.5%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNMKRV 11  
| | | | |  
Db 37 CLRQWQNMKRV 47

RESULT 7  
A84471  
En/Spm-like transposon protein [imported] - Arabidopsis thaliana  
C; Species: Arabidopsis thaliana (mouse-ear cress)  
C; Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C; Accession: A84471  
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999

A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A; Reference number: A84420; MUID: 20083487; PMID: 10617197  
A; Accession: A84471  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-531 <STO>  
A; Cross-references: GB: AE002093; NID: g4586022; PIDN: AAD25641.1; GSPDB: GN00139  
C; Genetics:  
A; Gene: At2g05650  
A; Map position: 2

Query Match 56.7%; Score 38; DB 2; Length 531;  
Best Local Similarity 75.0%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWXRNMK 10  
| | | | |  
Db 501 QWFRNMKX 508

RESULT 8  
T24218  
hypothetical protein R13G10.2 - Caenorhabditis elegans  
C; Species: Caenorhabditis elegans  
C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C; Accession: T24218  
R; Gardner, A.  
Submitted to the EMBL Data Library, August 1994  
A; Reference number: Z19857  
A; Accession: T24218  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-536 <WIL>  
A; Cross-references: EMBL: Z35602; PIDN: CAA84671.1; GSPDB: GN00021; CESP: R13G10.2  
A; Experimental source: clone R13G10  
C; Genetics:  
A; Gene: CESP: R13G10.2  
A; Map position: 3  
A; Introns: 64/3; 194/1; 404/3

Query Match 56.7%; Score 38; DB 2; Length 536;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
| | | | |  
Db 293 CIDWGRDDRKVK 304

RESULT 9  
T08030  
dynein beta heavy chain - Chlamydomonas reinhardtii  
C; Species: Chlamydomonas reinhardtii  
C; Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
C; Accession: T08030  
R; Mitchell, D.R.; Brown, K.S.  
J. Cell Sci. 107, 635-644, 1994  
A; Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
A; Reference number: Z16302; MUID: 94274778; PMID: 8006077  
A; Accession: T08030  
A; Status: translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-4568 <MIT>  
A; Cross-references: EMBL: U02963; NID: g409965; PIDN: AAA19956.1; PID: g514215  
A; Experimental source: strain 21gr  
C; Genetics:  
A; Gene: ODA4  
A; Map position: IX  
A; Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;  
3334/3; 3686/3; 3882/3; 4240/3  
C; Superfamily: dynein heavy chain, ciliary  
C; Keywords: nucleotide binding; P-loop  
F; 1919-1926/Region: nucleotide-binding motif A (P-loop)

F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 56.7%; Score 38; DB 2; Length 4568;  
Best Local Similarity 41.7%; Pred. No. 2.7e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 1852 CFQWQSLRVIQ 1863

#### RESULT 10

B60950

apolipoprotein B-100 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1993 #sequence\_revision 09-Sep-1994 #text\_change 07-Oct-1994

C:Accession: B60950

R:Law, A.; Scott, J.

J. Lipid Res. 31, 1109-1120, 1990

A:Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL  
A:Reference number: A60950; MUID:90324804; PMID:2373961

A:Accession: B60950

A:Molecule type: mRNA

A:Residues: 1-274 <LAW>

A:Note: authors translated the codon ATA for residue 8 as Val

C:Superfamily: apolipoprotein B

C:Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 55.2%; Score 37; DB 2; Length 274;

Best Local Similarity 66.7%; Pred. No. 27;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WXRNRKVR 12

Db 39 WDRNRKVR 47

#### RESULT 11

C70655

probable monooxygenase - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: C70655

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, C.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtrold, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70655

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-489 <COL>

A:Cross-references: GB:283864; GB:AL123456; NID:g3261687; PIDN:CAB06212.1; PID:e301250;

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: Rv3854c

Query Match 55.2%; Score 37; DB 2; Length 489;

Best Local Similarity 54.5%; Pred. No. 47;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 11

Db 253 CQWPRNRKRM 263

#### RESULT 12

T39801

hypothetical gh3-containing protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T39801

R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21880

A:Accession: T39801

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-501 <MCD>

A:Cross-references: EMBL:AL109731; PIDN:CAB52037.1; GSPDB:GN00067; SPDB:SPBCL9C2.10

A:Experimental source: strain 972h; cosmid c19C2

C:Genetics:

A:Gene: SPDB:SPBCL9C2.10

A:Map position: 2

A:Introns: 196/3

Query Match 55.2%; Score 37; DB 2; Length 501;

Best Local Similarity 60.0%; Pred. No. 48;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRK 10

Db 50 CTRWVRNMDK 59

#### RESULT 13

C84325

hypothetical protein Vngl732c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84325

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabi

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; I

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: C84325

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-584 <STO>

A:Cross-references: GB:AE004437; NID:g10581192; PIDN:AAG19967.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1732C

Query Match 55.2%; Score 37; DB 2; Length 584;

Best Local Similarity 41.7%; Pred. No. 56;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 445 CFTWRKDMERK 456

#### RESULT 14

C96582

F1511.22 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: C96582

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huiziar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96582

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-124 <STO>  
A;Cross-references: GB:AE005173; NID:g4587555; PIDN:AAD25786.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F1511.22  
A;Map position: 1

Query Match 53.7%; Score 36; DB 2; Length 124;  
Best Local Similarity 50.0%; Pred. No. 19;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 2 FQWXRNMKRV 11  
| | | | |  
Db 11 FRWGRSRREKI 20

RESULT 15  
AD2346  
hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp.  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C;Accession: AD2346  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; EMID:11759840  
A;Accession: AD2346  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-298 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BAE76022.1; PID:g17133459; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr4323

Query Match 53.7%; Score 36; DB 2; Length 298;  
Best Local Similarity 66.7%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 FQWXRNMKRV 10  
| | | | |  
Db 163 FHWQNVK 171

Search completed: February 21, 2003, 08:02:43  
Job time : 11.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107B-87  
Perfect score: 67  
Sequence: 1 CFQXRNMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	97.0	711	1 TRFL_HUMAN	P02788 homo sapien
2	48	71.6	708	1 TRFL_CAMDR	Q9tum0 camelus dro
3	48	71.6	708	1 TRFL_CAPHI	Q29477 capra hircu
4	40	59.7	292	1 NLA_DROME	Q9xz18 drosophila
5	40	59.7	695	1 TRFL_HORSE	Q07811 equus caball
6	39	58.2	707	1 TRFL_MOUSE	P08071 mus musculu
7	38	56.7	146	1 RPOB_LIBAF	P41187 liberibacte
8	38	56.7	783	1 YNR2_CABEL	Q21988 caenorhabdi
9	38	56.7	4568	1 DYHB_CHLRE	Q39565 chlamydomon
10	36	53.7	365	1 LA34_HUMAN	P30453 homo sapien
11	36	53.7	369	1 SPI1_MYXVL	P12393 myxoma viru
12	36	53.7	428	1 SYH_CHLMU	Q9pj19 chlamydia m
13	36	53.7	749	1 VF4_ROTGA	Q04916 rotavirus (
14	35	52.2	62	1 RL28_THETN	Q8r9ul thermoanaer
15	35	52.2	966	1 V1A_BBMV	Q00020 broad bean
16	35	52.2	1135	1 PHYC_SORBI	P93528 sorghum bic
17	34	50.7	214	1 VIF_SIV84	P12505 simian immu
18	34	50.7	422	1 PAF3_CHICK	Q30678 g platelet-
19	34	50.7	428	1 SYH_CHLFR	Q84547 chlamydia t
20	34	50.7	430	1 SYH_CHLFR	Q9z7p1 chlamydia p
21	34	50.7	455	1 YKYL_CABEL	Q19910 caenorhabdi
22	34	50.7	480	1 YQSL_CABEL	Q09309 caenorhabdi
23	34	50.7	500	1 TLCE_RICPR	Q05962 rickettsia
24	34	50.7	522	1 COR1_SCHPO	O13688 schizosacch
25	34	50.7	663	1 PDIL_HUMAN	Q9ulc6 homo sapien
26	34	50.7	765	1 Y008_HUMAN	Q15398 homo sapien
27	34	50.7	989	1 T100_HUMAN	O75448 homo sapien
28	34	50.7	1202	1 JAG2_RAT	P97607 rattus norv
29	34	50.7	1238	1 JAG2_HUMAN	Q9v219 homo sapien
30	34	50.7	1247	1 JAG2_MOUSE	Q9qy65 mus musculu
31	34	50.7	2151	1 RRPL_SEOUB	P27314 seoul virus
32	34	50.7	4568	1 DYHC_CABEL	Q19020 caenorhabdi
33	33.5	50.0	727	1 KDGA_RAT	P51556 rattus norv

## ALIGNMENTS

## RESULT 1

ID	TRFL_HUMAN	STANDARD;	PRT;	711 AA.
AC	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;			
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferronin A;			
DE	Lactoferronin B; Lactoferronin C].			
GN	LTF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rev M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cho Y.Y.;			
RA	Thesis (1994), Genetic Engineering Research Institute / Taejeon, Korea.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Connely O.N.;			
RL	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RL	"Molecular cloning and sequence analysis of human lactoferrin.";			
RN	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RL	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RL	sequences.";			
RN	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RL	"cDNA cloning and sequence analysis of human lactoferrin.";			
RN	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8]			

O88673 mus musculu  
Q99up4 staphylococ  
P16903 ovine lenti  
Q58118 methanococ  
P32174 escherichia  
P46421 arabidopsis  
Q9kdl8 bacillus ha  
Q55185 synechocyst  
P16710 vaccinia vi  
P20993 vaccinia vi  
P26929 lactobacill  
O95125 homo sapien

34 33.5 50.0 730 1 KDGA\_MOUSE  
35 33 49.3 62 1 RL28\_STRAAM  
36 33 49.3 144 1 REV\_OMVVS  
37 33 49.3 172 1 Y708\_METJA  
38 33 49.3 211 1 FDOI\_ECOLI  
39 33 49.3 224 1 GTXA\_ARATH  
40 33 49.3 238 1 CCDA\_BACHD  
41 33 49.3 246 1 Y495\_VSNY3  
42 33 49.3 275 1 VA16\_VACCV  
43 33 49.3 378 1 VA16\_VACCC  
44 33 49.3 573 1 URE1\_LACFE  
45 33 49.3 648 1 Z202\_HUMAN

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA";  
RL Nucleic Acids Res. 18:4013-4013 (1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RX Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins";  
RL Eur. J. Biochem. 145:659-666 (1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains";  
RL Biochim. Biophys. Acta 670:243-254 (1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin";  
RL FEBS Lett. 142:107-110 (1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis";  
RL Blood 70:989-993 (1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2595506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution";  
RL J. Mol. Biol. 209:711-734 (1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution";  
RL Acta Crystallogr. D 51:629-646 (1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253-->methionine mutant";  
RL Biochemistry 36:341-346 (1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori";  
RL Acta Crystallogr. D 55:403-407 (1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change";  
RL Acta Crystallogr. D 54:1319-1335 (1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin";  
RL Agric. Biol. Chem. 54:1803-1810 (1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene";  
RL Mol. Vision 4:31-32 (1998).  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -!- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS. WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X53961; CAA37914.1; -  
DR EMBL; U07643; AAB60324.1; -  
DR EMBL; M93150; AAA36159.1; -  
DR EMBL; M83202; AAA59511.1; -  
DR EMBL; M83205; AAA58656.1; -  
DR EMBL; M18642; AAA86665.1; -  
DR EMBL; AF332168; AAG48753.1; -  
DR EMBL; BC015822; AAH15822.1; -  
DR EMBL; BC015823; AAH15823.1; -  
DR EMBL; M73700; AAA59479.1; -  
DR EMBL; X52941; CAA37116.1; -  
DR EMBL; U95626; AAB57795.1; -  
DR PIR; S11228; TFHUL.  
DR PDB; 1LFC; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LFH; 31-OCT-93.  
DR PDB; 1LGI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1BKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.



Query Match 97.0%; Score 65; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00015;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB 39 CFQWXRNRKVR 50

## RESULT 2

TRFL\_CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9MZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin).  
 LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R.; Ackermann M.; Farah Z.; Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin."  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M.; Srinivasan A.; Singh R.; Sabani M.S.; Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AJ131674; CAB53387.1; -;  
 DR EMBL; AF165879; AAF82241.1; -;  
 DR HSSP; O77811; 1B1X.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 686 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 F -> S (IN REF. 2).  
 FT CONFLICT 261 261 G -> A (IN REF. 2).  
 FT CONFLICT 304 304 S -> P (IN REF. 2).  
 FT CONFLICT 330 330 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 492 494 L -> F (IN REF. 2).  
 FT CONFLICT 506 609 A -> P (IN REF. 2).  
 FT CONFLICT 609 609 R -> Q (IN REF. 2).  
 FT CONFLICT 642 642  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 71.6%; Score 48; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.21;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB 38 CAQWRMRKVR 49

## RESULT 3

TRFL\_CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F.; Nocart M.; Guerin G.; Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 RT relevant locus to bovine U12 syntenic group."  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.





Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

[4]  
 RN SEQUENCE OF 1-14 FROM N.A.  
 RP MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21880-21885(1991).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY CARBONATE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J03298; AAA40525.1; -;  
 CC EMBL; D88510; BAA13633.1; -;  
 CC EMBL; BC006904; AAH06904.1; -;  
 CC EMBL; M74778; AAA39427.1; -;  
 CC PIR; A28438; A28438.  
 CC HSP; P02788; 1CB6.  
 CC MGD; MGI:96837; ltf.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR\_FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN 1; 1.  
 CC PROSITE; PS00206; TRANSFERRIN 2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN 3; 2.  
 CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal.  
 KW SIGNAL.  
 FT CHAIN 1 19 BY SIMILARITY.  
 FT REPEAT 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.  
 FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 605 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139 ANION (POTENTIAL).  
 FT BINDING 461 461 ANION (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 MR -> IQG (IN REF. 1).  
 FT CONFLICT 1 2 R -> Q (IN REF. 2).  
 FT CONFLICT 25 25

FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;  
 Query Match 58.2%; Score 39; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 9.6;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWRNMRKV 11  
 DB 37 CLRQWEMRKV 47  
 ID RPOB LIBAF STANDARD; PRT; 146 AA.  
 AC P41187; 1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase  
 DE beta chain) RNA polymerase beta subunit (Fragment).  
 GN RPOB.  
 OS Liberibacter africanus (Liberibacter africanus).  
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Candidatus Liberibacter.  
 CX NCBI\_TaxID=34020;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nelspruit;  
 RA Planet P., Jagoueix S., Bove J.M., Garnier M.;  
 RT "Detection and characterization of the African Citrus Greening  
 RT Liberibacter by amplification, cloning and sequencing of the rplKAL-  
 RT rpoBC operon.";  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 CC BETA' CHAIN.  
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U09675; AAA19557.1; -;  
 CC InterPro; IPR001572; RNA\_pol\_B.  
 CC Pfam; PF00562; RNA\_pol\_B; 1.  
 CC PROSITE; PS01166; RNA POL BETA; PARTIAL.  
 KW Transferase; Transcription; DNA-directed RNA polymerase.  
 FT NON TER 146 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FD8943 CRC64;  
 Query Match 56.7%; Score 38; DB 1; Length 146;  
 Best Local Similarity 60.0%; Pred. No. 3;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWRNMRKV 10  
 DB 10 CVQWGRGARK 19

```

RESULT 8
YNR2 CAEEL STANDARD; PRT; 783 AA.
ID YNR2 CAEEL STANDARD; PRT; 783 AA.
AC Q2198;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein R13G10.2 in chromosome III.
GN R13G10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Gardner A.E.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RL
[2]
RN REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: FAD (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z35602; CAA84671.2; -
CC WormPep; R13G10.2; C525088.
CC InterPro; IPR002937; Amino oxidase.
CC Pfam; PF01593; Amino oxidase; 1.
CC Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
FT NP BIND 311 366 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 783 AA; 88799 MW; 8D087E96464DC908 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 783;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWXRNMRKVR 12
DB 540 CIDWGRDRKVK 551

RESULT 9
DYHB CHLRE STANDARD; PRT; 4568 AA.
ID DYHB CHLRE STANDARD; PRT; 4568 AA.
AC Q39565;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dynein beta chain, flagellar outer arm.
GN ODA4 OR ODA-4 OR SUP1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=2197;
RC Mitchell D.R.; Brown K.S.;
RA MEDLINE=94274778; PubMed=8006077;
RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy
RT chain genes."
RL J. Cell Sci. 107:635-644 (1994).
CC -!- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND
CC FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.

```

```

CC DYNEIN HAS ATPASE ACTIVITY.
CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND
CC GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.
CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02963; AAL19556.1; -
CC InterPro; IPR004273; Dynein heavy.
CC Pfam; PF03028; Dynein heavy; 1.
CC Motor protein; Microtubules; Dynein; ATP-binding; Flagella;
KW Coiled coil.
FT DOMAIN 277 293 COILED COIL (POTENTIAL).
FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).
FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).
FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).
FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).
FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).
FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).
FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).
FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).
FT NP BIND 1919 1926 ATP (POTENTIAL).
FT NP BIND 2202 2209 ATP (POTENTIAL).
FT NP BIND 2530 2537 ATP (POTENTIAL).
FT NP BIND 2879 2886 ATP (POTENTIAL).
SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 4568;
Best Local Similarity 41.7%; Pred. No. 95;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWXRNMRKVR 12
DB 1852 CFOWSQLRYIQ 1863

RESULT 10
IA34 HUMAN STANDARD; PRT; 365 AA.
ID IA34 HUMAN STANDARD; PRT; 365 AA.
AC P30453; P30454;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HLA class I histocompatibility antigen, A*34(A-10) alpha chain
DE precursor.
GN HLA-A OR HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=93056508; PubMed=1431115;
RA Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA Martell R.W., du Toit E.D., Parham P.;
RT "Distinctive HLA-A,B antigens of black populations formed by
RT interallelic conversion."
RL J. Immunol. 149:3411-3415 (1992).
[2]
RN SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=93235211; PubMed=8475492;
RA Madrigal J.A., Hildebrand W.H., Belich M.P., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA du Toit E.D., Parham P.;
RT "Structural diversity in the HLA-A10 family of alleles: correlations

```

RT with serology.";  
 RL Tissue Antigens 41:72-80(1993).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -!- POLYMORPHISM: THE FOLLOWING ALLELES OF AW-34 ARE KNOWN: A\*3401  
 CC (AW-34.1) AND A\*3402 (AW34.2). THE SEQUENCE SHOWN IS THAT OF  
 CC A\*3401.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X61704; CAA43873.1; -;  
 DR EMBL; X61705; CAA43874.1; -;  
 DR PIR; S16767; S16767.  
 DR PIR; S16771; S16771.  
 DR HSSP; O19673; IHSB.  
 DR MIM; 142800; -;  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR001039; MHC\_I.  
 DR Pfam; PF00047; Ig\_1.  
 DR Pfam; PF00129; MHC\_I\_1.  
 DR ProDom; PD000050; MHC\_I; 1.  
 DR SMART; SM00407; IgC1; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.  
 FT SIGNAL 1 24  
 FT CHAIN 25 365  
 FT -----  
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT AW-34 (A-10) ALPHA CHAIN.  
 FT EXTRACELLULAR ALPHA-1.  
 FT EXTRACELLULAR ALPHA-2.  
 FT EXTRACELLULAR ALPHA-3.  
 FT CONNECTING PEPTIDE.  
 FT -----  
 FT CYTOPLASMIC TAIL.  
 FT N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
 FT BY SIMILARITY.  
 FT BY SIMILARITY.  
 FT .I -> V (IN A\*3402).  
 FT /FTID=VAR 004379.  
 FT K -> N (IN A\*3402).  
 FT /FTID=VAR 004380.  
 FT R -> I (IN A\*3402).  
 FT /FTID=VAR 004381.  
 FT P -> S (IN A\*3402).  
 FT /FTID=VAR 004382.  
 FT Q -> R (IN A\*3402).  
 FT /FTID=VAR 004383.  
 FT W -> L (IN A\*3402).  
 FT /FTID=VAR 004384.  
 FT L -> I (IN A\*3402).  
 FT /FTID=VAR 004385.  
 FT -----  
 SQ SEQUENCE 365 AA; 41055 MW; 063BF63E6E01F6 CRC64;  
 Query Match 53.7%; Score 36; DB 1; Length 365;  
 Best Local Similarity 66.7%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 WXRNRKVR 12  
 | | | | |  
 Db 84 WDRNTRKVK 92  
 RESULT 11  
 SP11\_MXVVL STANDARD; PRT; 369 AA.  
 ID SP11\_MXVVL

AC P12393;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serine proteinase inhibitor 1 (Serp-1).  
 GN SERP1 OR M008.1L.  
 OS Myxoma virus (strain Lausanne).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Leporipoxvirus.  
 OX NCBI\_TaxID=31530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91049428; PubMed=2173255;  
 RA Upton C., Macen J.L., Wishart D.S., McFadden G.;  
 RT "Myxoma virus and malignant rabbit fibroma virus encode a serpin-like  
 RT protein important for virus virulence.";  
 RL Virology 179:618-631(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87064296; PubMed=3023828;  
 RA Upton C., McFadden G.;  
 RT "DNA sequence homology between the terminal inverted repeats of Shope  
 RT fibroma virus and an endogenous cellular plasmid species.";  
 RL Mol. Cell. Biol. 6:265-276(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87030884; PubMed=3021526;  
 RA Upton C., Carrell R.W., McFadden G.;  
 RT "A novel member of the serpin superfamily is encoded on a circular  
 RT plasmid-like DNA species isolated from rabbit cells.";  
 RL FEBS Lett. 207:115-120(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20032073; PubMed=10562494;  
 RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,  
 RA Macaulay C., Willer D., Evans D., McFadden G.;  
 RT "The complete DNA sequence of myxoma virus.";  
 RL Virology 264:298-318(1999).  
 CC -!- FUNCTION: IMPORTANT IN VIRULENCE.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. POXVIRUSES SUBFAMILY.  
 CC -!- CAUTION: WAS ORIGINALLY (REF.2 AND REF.3) THOUGHT TO ORIGINATE  
 CC FROM A PLASMID RABBIT DNA. THE ORIGINAL SAMPLE WAS CONTAMINATED  
 CC AND THE GENE IS DERIVED FROM MYXOMA VIRUS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M35233; AAA46629.1; -;  
 DR EMBL; M12333; AAA81567.1; -;  
 DR EMBL; AF170726; AAF15055.1; -;  
 DR EMBL; AF170726; AAF14896.1; -;  
 DR PIR; A24470; A24470.  
 DR PIR; B36418; B36418.  
 DR HSSP; P05121; IB3K.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Virulence.  
 FT ACT SITE 319 320 REACTIVE BOND (BY SIMILARITY).  
 FT CONFLICT 1 16 MKYLVVLCLTSACR -> MFNVVRV (IN REF. 2  
 FT AND 3).  
 FT SEQUENCE 369 AA; 41556 MW; 8DB31CE131C218A0 CRC64;  
 Query Match 53.7%; Score 36; DB 1; Length 369;  
 Best Local Similarity 85.7%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WYRNWK 10  
 DB 235 WYRNWK 241

RESULT 12  
 ID SYH CHLMU STANDARD; PRT; 428 AA.  
 AC Q9PJU9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)  
 DE (HisRS).  
 GN HISS OR TC0830.  
 OS Chlamydia muridarum.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MoPn / Nig9;  
 RX MEDLINE=20150255; PubMed=10694935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 White O., Hickman E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 Eisen J., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 Eisen J., Fraser C.M.,  
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
 pneumoniae AR39.",  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +  
 diphosphate + L-histidyl-tRNA(His).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M91434; AAA47338.1;  
 DR InterPro; IPR000416; Cap\_VP4.  
 DR Pfam; PF00426; VP4; 1.  
 KW Coat protein; Glycoprotein.  
 KM CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 681 681 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 698 698 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 749 AA; 84362 MW; D123527DEAE0F21 CRC64;

Query Match 53.7%; Score 36; DB 1; Length 749;  
 Best Local Similarity 50.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 DB 195 CFTWDMNCANVR 206

RESULT 14  
 ID RL28\_THETN STANDARD; PRT; 62 AA.  
 AC Q8R9U1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE SOS ribosomal protein L28.  
 GN RPMB OR TTE1495.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao O., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 Tan H., Chen R., Wang J., Yu J., Yang H.,  
 RT "A complete sequence of T. tengcongensis genome."  
 RL Genome Res. 12:689-700(2002).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.

QY 4 WYRNWK 10  
 DB 235 WYRNWK 241

RESULT 13  
 ID VP4\_ROTGA STANDARD; PRT; 749 AA.  
 AC Q04916;  
 DT 01-JUN-1994 (Rel. 29, Created)

Query Match 53.7%; Score 36; DB 1; Length 428;  
 Best Local Similarity 44.4%; Pred. No. 21;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNR 9  
 DB 348 CFSWAKHLR 356

Db 347 CFKKNKNTNRSV 361  
Search completed: February 21, 2003, 07:51:35  
Job time : 6.2093 secs

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: AE013107; AAM24713.1; -  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match 52.2%; Score 35; DB 1; Length 62;  
Best Local Similarity 60.0%; Pred. No. 4.6;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 QWXRNRKVR 12  
Db 27 RWKPNIRKVR 36

## RESULT 15

VIA\_BMV VIA\_BMV STANDARD; PRT; 966 AA.  
AC Q00020;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1A protein [Includes: Helicase; Methyltransferase].  
OS Broad bean mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;  
OC Bromovirus.  
OX NCBI\_TaxID=12301;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bawden;  
RX MEDLINE=92074218; PubMed=1962437;  
RA Dianott A.M., Bujarski J.J.;  
RT "The nucleotide sequence and genome organization of the RNA-1 segment  
RT in two bromoviruses: broad bean mottle virus and cowpea chlorotic  
RT mottle virus.";  
RL Virology 185:553-562(1991).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.  
CC -!- CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE  
CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA  
CC CAPPING.  
CC -!- SIMILARITY: TO 1A PROTEIN FROM CCMV, CMV, PSV AND TAV.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: M65138; AAA42740.1; -  
DR PIR: A41699; PIBVBB.  
DR InterPro: IPR002588; V\_methyltransf.  
DR InterPro: IPR000606; Viral\_helicase1.  
DR Pfam: PF01443; Viral\_helicase1.  
DR Pfam: PF01660; Vmethyltransf; 1.  
KW Helicase; ATP-binding; Transferase; Methyltransferase.  
FT NP BIND 690 697 ATP (POTENTIAL).  
SQ SEQUENCE 966 AA; 109621 MW; DF592681D7231C8D CRC64;

Query Match 52.2%; Score 35; DB 1; Length 966;  
Best Local Similarity 46.7%; Pred. No. 72;  
Matches 7; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

OY 1 CFQ----WXRNRKVR 11  
|||



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-87

Perfect score: 67

Sequence: 1 CFQWXRMRKV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_podent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	89.6	711	4	Q8TCD2
2	56	83.6	38	4	Q9UCY5
3	45	67.2	33	6	Q9TR80
4	41	61.2	511	16	Q8Z462
5	40	59.7	148	10	Q9XHP1
6	40	59.7	273	2	Q31090
7	39	58.2	275	5	Q93780
8	39	58.2	341	11	Q8R2A4
9	39	58.2	415	5	Q9U4M9
10	39	58.2	866	10	Q9FH19
11	38	56.7	108	9	Q8SC55
12	38	56.7	372	10	Q81653
13	38	56.7	531	10	Q9SIC0
14	37	55.2	91	15	Q77855
15	37	55.2	91	15	Q77856
16	37	55.2	179	11	Q9CZQ7

17	37	55.2	233	11	Q9D4X5
18	37	55.2	274	4	Q96M21
19	37	55.2	279	16	Q8XSE2
20	37	55.2	323	12	Q9QB73
21	37	55.2	368	11	Q9DAJ3
22	37	55.2	368	11	Q9D9X0
23	37	55.2	368	11	Q9JIT1
24	37	55.2	377	12	Q91MQ5
25	37	55.2	381	12	Q9DHK5
26	37	55.2	489	16	P96223
27	37	55.2	501	3	Q9UUD0
28	37	55.2	570	10	Q8S487
29	37	55.2	584	17	Q9HPA3
30	37	55.2	2348	5	Q9V346
31	36	53.7	105	10	Q9XFD5
32	36	53.7	109	15	Q9YQC1
33	36	53.7	109	15	Q9YQC0
34	36	53.7	109	15	Q9YQB9
35	36	53.7	109	15	Q9YQB8
36	36	53.7	109	15	Q9YJ17
37	36	53.7	109	15	Q9YJ12
38	36	53.7	119	15	Q8Q454
39	36	53.7	124	10	Q9SYH0
40	36	53.7	298	16	Q8YP77
41	36	53.7	306	4	Q8TAX2
42	36	53.7	466	4	Q9NUS2
43	36	53.7	514	5	Q9V118
44	36	53.7	514	10	Q9SP27
45	36	53.7	543	5	Q9XZ30

## ALIGNMENTS

### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAB22347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 89.6%; Score 60; DB 4; Length 711;

Best Local Similarity 90.9%; Pred. No. 0.003;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRMRKV 11

DB 39 CFQWXRMRKV 49

### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.

AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE Lactoferrin homolog (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96081613; PubMed=8551695;
RA Sato I.;
RT "Characterization of the 84-kDa protein with ABH activity in human
  seminal plasma.";
RL Jpn. J. Legal Med. 49:281-293 (1995).
DR HSP; P02788; 1BKA.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDD8B CRC64;

Query Match 83.6%; Score 56; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.00085;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWXRNRKVR 12
DB 21 FQWXRNRKVR 31

RESULT 3
Q9TR80 PRELIMINARY; PRT; 33 AA.
AC Q9TR80;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Lactoferrin (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=95127729; PubMed=7827104;
RA Qian Z.Y., Jolles P., Migliore-Samouir D., Fiat A.M.;
RL Biochim. Biophys. Acta 1243:25-32 (1995).
DR HSP; O77638; 1CE2.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 1.
SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 67.2%; Score 45; DB 6; Length 33;
Best Local Similarity 54.5%; Pred. No. 0.098;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11
DB 19 CYQWQKMRKL 29

RESULT 4
Q8Z462 PRELIMINARY; PRT; 511 AA.
AC Q8Z462;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein STY3070.
GN STY3070.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
"Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
DR EMBL; AL627276; CAD06049.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 511 AA; 58126 MW; E2DD124E10D178B CRC64;

Query Match 61.2%; Score 41; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 9.9;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
DB 350 CFQWXRNRKVR 361

RESULT 5
Q9XHP1 PRELIMINARY; PRT; 148 AA.
AC Q9XHP1;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE 2S albumin.
OS Sesamum indicum (Oriental sesame) (Gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TAINAN 1;
RX MEDLINE=20074970; PubMed=10606554;
RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;
RT "Molecular cloning of 11S globulin and 2S albumin, the two major seed
  storage proteins in sesame.";
RL Agric. Food Chem. 47:4932-4938 (1999).
DR EMBL; AF01841; XAD42943.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR00617; Napin.
DR Pfam; PF00234; trypan alpha_ami; 1.
DR PRINTS; PR00496; NAFIN.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 59.7%; Score 40; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 4.3;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRMR 9.
DB 54 CMQWXRMR 62

RESULT 6
Q31090 PRELIMINARY; PRT; 273 AA.
AC Q31090;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE Hypothetical 31.0 kDa protein.
OS Rhizobium leguminosarum (biovar viciae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
```

```

OX NCBI_TaxID=387;
RN SEQUENCE FROM N.A.
RP STRAIN=VF39;
RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=VF39;
RX MEDLINE=99113394; PubMed=9914965;
RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,
RL Ksenzenko V.N.;
RT "Structural and functional organization of the exopolysaccharide
RT biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39.";
RL Mol. Biol. (Mosk) 32:797-804(1998).
DR EMBL; AF028810; AAB8891.1; -.
KW Hypothetical protein.
SQ SEQUENCE 273 AA; 30986 MW; F195D2FEF7AD4D3 CRC64;

Query Match 59.7%; Score 40; DB 2; Length 273;
Best Local Similarity 60.0%; Pred. No. 8.1;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QWXRNRKVR 12
Db 245 RWLRNLRKLR 254

RESULT 7
Q93780
ID Q93780 PRELIMINARY; PRT; 275 AA.
AC Q93780;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 13, Last annotation update)
DE F53H4.4 protein.
GN F53H4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN SEQUENCE FROM N.A.
RA Dobson R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81089; CAB03137.1; -.
SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 58.2%; Score 39; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FQWXRNRKVR 12
Db 262 FQWXRNRKVR 272

RESULT 8
Q8R2A4
ID Q8R2A4 PRELIMINARY; PRT; 341 AA.
AC Q8R2A4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vomeronasal receptor VIREG.
GN VIREG.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP STRAIN=129X1/SVJ;
RX MEDLINE=21676859; PubMed=11802169;
RA Rodriguez I., Del Punta K., Rothman A., Ishii T., Mombaerts P.;
RT "Multiple new and isolated families within the mouse superfamily of
RT V1r vomeronasal receptors.";
RL Nat. Neurosci. 5:134-140(2002).
DR EMBL; AY065506; AAL47911.1; -.
KW Receptor.
SQ SEQUENCE 341 AA; 39632 MW; B3D8259F7AD11130 CRC64;

Query Match 58.2%; Score 39; DB 11; Length 341;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNR 9
Db 331 CFTWTRNIK 339

RESULT 9
Q9U4M9
ID Q9U4M9 PRELIMINARY; PRT; 415 AA.
AC Q9U4M9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 7138.2.
GN 7138.2.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN SEQUENCE FROM N.A.
RP STRAIN=FRIEDLIN;
RA Ravel C., Dubessay P., Blaineau C., Lignon M.-F., Bastien P.,
RA Dedet J.-P., Pages M.;
RT "Leishmania major chromosome 5 complete sequence.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163772; AAF14642.1; -.
DR InterPro; IPR003409; MORN.
DR Pfam; PF02493; MORN; 4.
SQ SEQUENCE 415 AA; 46701 MW; 3E3AD710BF23691E CRC64;

Query Match 58.2%; Score 39; DB 5; Length 415;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNR 9
Db 272 CGEWFNRNR 280

RESULT 10
Q9FHI9
ID Q9FHI9 PRELIMINARY; PRT; 866 AA.
AC Q9FHI9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genomic DNA, chromosome 5, pl clone:MFC19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.

```

RC STRAIN=COLUMBIA;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT P1 and TAC clones.";  
 RL DNA Res. 6:183-195(1999).  
 DR EMBL; AB018113; BAB0175.1; -;  
 SQ SEQUENCE 866 AA; 100187 MW; 164FAB9B4976D27D CRC64;

Query Match 58.2%; Score 39; DB 10; Length 866;  
 Best Local Similarity 50.0%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||:|  
 DB 389 CFNWLKFKRLR 400

RESULT 11  
 Q8SC55  
 ID Q8SC55 PRELIMINARY; PRT; 108 AA.  
 AC Q8SC55;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21; Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21; Last annotation update)  
 DE Hypothetical 12.3 kDa protein.  
 OS Stx2 converting bacteriophage I.  
 OC Viruses.  
 OX NCBI\_TaxID=180816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=STX2 PHAGE-I;  
 RA Sato T., Shimizu T., Watarai M., Kobayashi M., Kano S., Hamabata T.,  
 RA Yamaeaki S., Takeda Y.;  
 RT "Genomic sequence of Shiga toxin 2-converting phage isolated from  
 RT Escherichia coli O157:H7 Okayama strain and comparison with other  
 RT Shiga toxin 2-converting phages.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF004402; BAB87947.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 108 AA; 12279 MW; 2FF2E76690C78BA7 CRC64;

Query Match 56.7%; Score 38; DB 9; Length 108;  
 Best Local Similarity 60.0%; Pred. No. 7.5;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRK 10  
 |||:|  
 DB 31 CFQWISNRK 40

RESULT 12  
 O81653  
 ID O81653 PRELIMINARY; PRT; 372 AA.  
 AC O81653;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08; Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20; Last annotation update)  
 DE Senescence-associated protein 3 (Fragment).  
 GN SA3.  
 OS Hemerocallis hybrid cultivar.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
 OC Hemerocallidaceae; Hemerocallis.  
 OX NCBI\_TaxID=80862;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. STELLA D'ORO; TISSUE=SENESCING PETALS;  
 RX MEDLINE=99339248; PubMed=10412903;  
 RA Panavas T., Pikula A., Reid P.D., Rubinstein B., Walker E.L.;  
 RT "Identification of senescence-associated genes from daylily petals.";

RL Plant Mol. Biol. 40:237-248(1999).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF082028; AAC34853.1; -;  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PRO0385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 FT NON\_TER 1  
 SQ SEQUENCE 372 AA; 41909 MW; 55FE3EAF9699595E CRC64;

Query Match 56.7%; Score 38; DB 10; Length 372;  
 Best Local Similarity 71.4%; Pred. No. 27;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRN 7  
 |||:|  
 DB 328 CFWDEN 334

RESULT 13  
 Q9SIC0  
 ID Q9SIC0 PRELIMINARY; PRT; 531 AA.  
 AC Q9SIC0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13; Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19; Last annotation update)  
 DE En/Spm-like transposon protein.  
 GN AT2G05650.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Renito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.B., Barnstead M.E., Feldblum T.V.,  
 RA Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.B.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC007170; AAD25641.1; -;  
 DR InterPro; IPR004264; Transposase 23.  
 DR Pfam; PF03017; Transposase 23; 1.  
 SQ SEQUENCE 531 AA; 60512 MW; 57B3AC60C976A4B9 CRC64;

Query Match 56.7%; Score 38; DB 10; Length 531;  
 Best Local Similarity 75.0%; Pred. No. 39;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWXRNRK 10  
 |||:|  
 DB 501 QWFRNRK 508

RESULT 14  
 Q77855  
 ID Q77855 PRELIMINARY; PRT; 91 AA.  
 AC Q77855;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01; Last sequence update)

DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE Envelope protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95191002; PubMed=7884875;  
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,  
RT Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;  
RT "Similarity in env and gag genes between genomic RNAs of human  
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is  
RT unrelated to time of HIV-1 RNA positivity in the child.";  
RL J. Virol. 69:2285-2296(1995).  
DR EMBL; Z47867; CAA87881.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 1  
FT NON\_TER 91  
SQ SEQUENCE 91 AA; 10530 MW; 8B10C62011F305D6 CRC64;  
Query Match 55.2%; Score 37; DB 15; Length 91;  
Best Local Similarity 60.0%; Pred. No. 9.9;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 QWXRNMKVR 12  
Db ||| :|||  
67 QWRTLOKVR 76

RESULT 15  
Q77856 PRELIMINARY; PRT; 91 AA.  
AC Q77856;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE Envelope protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95191002; PubMed=7884875;  
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,  
RT Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;  
RT "Similarity in env and gag genes between genomic RNAs of human  
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is  
RT unrelated to time of HIV-1 RNA positivity in the child.";  
RL J. Virol. 69:2285-2296(1995).  
DR EMBL; Z47868; CAA87882.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Glycoprotein.  
FT NON\_TER 1  
FT NON\_TER 1  
FT NON\_TER 91  
SQ SEQUENCE 91 AA; 10519 MW; EB20C63A22DA1288 CRC64;  
Query Match 55.2%; Score 37; DB 15; Length 91;  
Best Local Similarity 60.0%; Pred. No. 9.9;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 QWXRNMKVR 12  
Db ||| :|||  
67 QWRTLOKVR 76

Search completed: February 21, 2003, 08:00:42  
Job time : 21.6744 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-88

Perfect score: 67

Sequence: 1 CFQWXRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	98.5	12	21 AAY78072	Human lactoferrin
2	65	97.0	12	21 AAY78038	Human lactoferrin
3	65	97.0	12	21 AAY78046	Human lactoferrin
4	65	97.0	12	21 AAY78047	Human lactoferrin
5	65	97.0	12	21 AAY78084	Human lactoferrin
6	65	97.0	12	21 AAY78087	Human lactoferrin
7	65	97.0	12	21 AAY78088	Human lactoferrin
8	65	97.0	12	21 AAY78091	Human lactoferrin
9	65	97.0	12	21 AAY78092	Human lactoferrin
10	65	97.0	13	21 AAY78037	Human lactoferrin

11	65	97.0	13	21 AAY78048	Human lactoferrin
12	65	97.0	13	21 AAY78049	Human lactoferrin
13	65	97.0	14	21 AAY78036	Human lactoferrin
14	65	97.0	14	21 AAY78050	Human lactoferrin
15	65	97.0	14	21 AAY78051	Human lactoferrin
16	65	97.0	15	17 AAR98554	Peptide for anti-u
17	65	97.0	15	21 AAY78035	Human lactoferrin
18	65	97.0	15	21 AAY78062	Human lactoferrin
19	65	97.0	15	21 AAY78063	Human lactoferrin
20	65	97.0	16	21 AAY78031	Human lactoferrin
21	65	97.0	16	21 AAY78064	Human lactoferrin
22	65	97.0	16	21 AAY78065	Human lactoferrin
23	65	97.0	17	21 AAY78034	Human lactoferrin
24	65	97.0	17	21 AAY78066	Human lactoferrin
25	65	97.0	17	21 AAY78067	Human lactoferrin
26	65	97.0	18	15 AAR69352	Human lactoferrin
27	65	97.0	18	17 AAW13397	Human lactoferrin
28	65	97.0	18	21 AAY78033	Human lactoferrin
29	65	97.0	19	21 AAY68867	Amino acid sequenc
30	65	97.0	19	21 AAY78032	Human lactoferrin
31	65	97.0	20	13 AAR21810	Anti microbial pep
32	65	97.0	20	14 AAR44841	Lactoferrin-relate
33	65	97.0	20	15 AAR48530	Lactoferrin-derive
34	65	97.0	20	15 AAR48531	Lactoferrin-derive
35	65	97.0	20	15 AAR57461	Lactoferrin-derive
36	65	97.0	20	15 AAR57462	Lactoferrin-derive
37	65	97.0	20	16 AAR84698	Bovine lactoferrin
38	65	97.0	20	16 AAR84699	Bovine lactoferrin
39	65	97.0	20	16 AAR80263	Anti-parasitic lac
40	65	97.0	20	16 AAR80264	Anti-parasitic lac
41	65	97.0	20	17 AAR98553	Peptide for anti-u
42	65	97.0	20	17 AAR91852	Lactoferrin-derive
43	65	97.0	20	17 AAW03045	Lactoferrin-derive
44	65	97.0	20	17 AAR90607	Lactoferrin-derive
45	65	97.0	20	17 AAR87621	Lactoferrin-derive

#### ALIGNMENTS

#### RESULT 1

AAV78072  
ID AAV78072 standard; Peptide; 12 AA.  
XX  
AC AAY78072;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:72.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 22; Page 35; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 3.1e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 12  
 Db 1 CFQWXRNRKVR 12

RESULT 2

AAY78038  
 ID AAY78038 standard; Peptide; 12 AA.

XX AC AAY78038;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:38.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX FN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 70; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 12  
 Db 1 CFQWXRNRKVR 12

RESULT 3

AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.

XX AC AAY78046;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:46.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX FN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 15; Page 35; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWQRNRKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:47.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 73; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, colitis, and Candida infection on a mucosal  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWQRNRKVR 12

RESULT 5  
 AAY78084

ID AAY78084 standard; Peptide; 12 AA.

XX AC AAY78084;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:84.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations, colitis, and Candida infection on a mucosal  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 1 CFQWXRNRKVR 12  
 |||||  
 Db 1 CFQWXRNRKVR 12

## RESULT 6

AAAY78087  
 ID AAY78087 standard; Peptide; 12 AA.

AC AAY78087;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:87.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.7e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 1 CFQWXRNRKVR 12

## RESULT 7

AAAY78088

ID AAY78088 standard; Peptide; 12 AA.

XX AAY78088;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:88.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

PI WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 65; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.7e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 1 CFQWXRNRKVR 12

## RESULT 8

AAAY78091

ID AAY78091 standard; Peptide; 12 AA.

XX AAY78091;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:91.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SR01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 XX Claim 22; Page 38; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWKXNMRKVR 12  
 Db |||||  
 1 CFQWKXNMRKVR 12  
 RESULT 9  
 AAY78092  
 ID AAY78092 standard; Peptide; 12 AA.  
 XX  
 AC AAY78092;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 DE Human lactoferrin derived peptide SEQ ID NO:92.  
 XX  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 DN  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SR01230.  
 PF  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 XX Claim 22; Page 38; 102pp; English.  
 PS  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 97.0%; Score 65; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWKXNMRKVR 12  
 Db |||||  
 1 CFQWKXNMRKVR 12  
 RESULT 10  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.  
 XX  
 AC AAY78037;  
 XX  
 XX 25-APR-2000 (first entry)  
 DT  
 DE Human lactoferrin derived peptide SEQ ID NO:37.  
 XX  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200001730-A1.  
 PN  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX  
 XX 06-JUL-1999; 99WO-SR01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 70; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 13 AA;  
 SQ Query Match 97.0%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 5e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKVR 12  
 DB ||||| |||||  
 2 CFQWQRNRKVR 13  
 RESULT 11  
 AAY78048  
 ID AAY78048 standard; Peptide; 13 AA.  
 XX AC AAY78048;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:48.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 74; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 13 AA;  
 SQ Query Match 97.0%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 5e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKVR 12  
 DB ||||| |||||  
 2 CFQWQRNRKVR 13  
 RESULT 12  
 AAY78049  
 ID AAY78049 standard; Peptide; 13 AA.  
 XX AC AAY78049;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:49.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PS Claim 18; Page 74; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 97.0%; Score 65; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 5e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 ||||| |||||  
 Db 2 CFQWQRNRKVR 13

RESULT 13

AAV78036  
 ID AAY78036 standard; Peptide; 14 AA.

XX AC AAY78036;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:36.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,

XX PT inflammations and tumors and for use in infant formula food -

XX XX Claim 12; Page 69; 102pp; English.

XX PS AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
 ||||| |||||  
 Db 3 CFQWQRNRKVR 14

RESULT 14

AAV78050  
 ID AAY78050 standard; Peptide; 14 AA.

XX AC AAY78050;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:50.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCII-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,

XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 75; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides can also be used  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 5.4e-05;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 3 CFQWQRNRKVR 14

RESULT 15

AAAY78051

ID AAY78051 standard; Peptide; 14 AA.

AC AAY78051;

DT 25-APR-2000 (First entry)

XX XX

DE Human lactoferrin derived peptide SEQ ID NO:51.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 75; 102pp; English.

PS AAY78001 to AAY78100 represent peptides having sequences based on human

XX lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumors. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 65; DB 21; Length 14;

Best Local Similarity 91.7%; Pred. No. 5.4e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12

Db 3 CFQWQRNRKVR 14

Search completed: February 21, 2003, 07:56:43

Job time : 29.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-88

Perfect score: 67

Sequence: 1 CFQXNRMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PGTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	97.0	18	1	US-08-204-487-3
2	65	97.0	18	2	US-08-485-948-8
3	65	97.0	18	2	US-08-628-380-8
4	65	97.0	18	2	US-08-475-055-8
5	65	97.0	20	1	US-07-755-161A-3
6	65	97.0	20	1	US-07-891-174-3
7	65	97.0	20	1	US-08-204-487-1
8	65	97.0	20	1	US-08-256-771-24
9	65	97.0	20	1	US-08-256-771-25
10	65	97.0	20	1	US-08-381-984-24
11	65	97.0	20	1	US-08-381-984-25
12	65	97.0	22	4	US-09-508-734-4
13	65	97.0	24	4	US-09-508-734-6
14	65	97.0	25	1	US-07-755-161A-10
15	65	97.0	25	1	US-07-891-174-10
16	65	97.0	25	1	US-08-204-487-7
17	65	97.0	29	4	US-09-508-734-8
18	65	97.0	36	1	US-07-755-161A-8
19	65	97.0	36	1	US-07-891-174-8
20	65	97.0	36	1	US-08-256-771-30
21	65	97.0	36	1	US-08-381-984-29
22	65	97.0	47	2	US-08-464-182A-6
23	65	97.0	47	2	US-08-406-271-6
24	65	97.0	50	2	US-08-693-274A-7
25	65	97.0	52	4	US-09-017-043A-3
26	65	97.0	53	2	US-08-464-182A-5
27	65	97.0	53	2	US-08-406-271-5

28	65	97.0	54	2	US-08-464-182A-2	Sequence 2, Appli
29	65	97.0	54	2	US-08-406-271-2	Sequence 2, Appli
30	65	97.0	694	3	US-08-724-586-2	Sequence 2, Appli
31	65	97.0	694	4	US-09-421-632-2	Sequence 2, Appli
32	65	97.0	694	4	US-09-932-190-2	Sequence 2, Appli
33	65	97.0	705	4	US-08-655-640-2	Sequence 2, Appli
34	65	97.0	708	2	US-08-655-640-4	Sequence 2, Appli
35	65	97.0	711	1	US-08-154-019-4	Sequence 4, Appli
36	65	97.0	711	1	US-08-461-333-4	Sequence 4, Appli
37	65	97.0	711	3	US-08-464-167-4	Sequence 4, Appli
38	65	97.0	711	3	US-09-158-313-4	Sequence 4, Appli
39	65	97.0	711	4	US-08-476-798-4	Sequence 4, Appli
40	62	92.5	711	1	US-08-145-681-2	Sequence 2, Appli
41	62	92.5	711	1	US-08-250-308-2	Sequence 2, Appli
42	62	92.5	711	1	US-08-453-703-2	Sequence 2, Appli
43	62	92.5	711	2	US-08-456-106-2	Sequence 2, Appli
44	62	92.5	711	3	US-08-456-108-2	Sequence 2, Appli
45	62	92.5	711	4	US-09-265-577-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5563425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 97.0%; Score 65; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5900  
; TELEFAX: 201 343-1684  
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 97.0%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 97.0%; Score 65; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.6e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
||| |||||  
Db 1 CFQWRNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI

; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/489,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LF-C1, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

Query Match 97.0%; Score 65; DB 2; Length 18;
Best Local Similarity 91.7%; Pred No. 3.6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNRKVR 12
Db 1 CFQWRNRKVR 12

RESULT 5
US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible
```

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3
```

Query Match

97.0%; Score 65; DB 1; Length 20;



Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRNRKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRNRKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNNRKVR 12  
||| |||||  
Db 2 CFQWRNNRKVR 13

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"

US-08-256-771-24

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNNRKVR 12  
||| |||||  
Db 2 CFQWRNNRKVR 13

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
US-08-256-771-25

Query Match 97.0%; Score 65; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNNRKVR 12  
||| |||||  
Db 2 CFQWRNNRKVR 13

RESULT 10  
US-08-381-984-24

```

; Sequence 24, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof
; US-08-381-984-24

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
DB 2 CFQWRNRKVR 13

RESULT 11
US-08-381-984-25
; Sequence 25, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof
; US-08-381-984-24

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
DB 2 CFQWRNRKVR 13

RESULT 11
US-08-381-984-25
; Sequence 25, Application US/08381984
; Patent No. 5804555
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIOXIDANT
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"
; US-08-381-984-25

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
DB 2 CFQWRNRKVR 13

RESULT 12
US-09-508-734-4
; Sequence 4, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; TITLE OF INVENTION: useful microorganism thereof
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KE99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KRI998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 4

```

```

; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,984
; FILING DATE: April 11, 1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "the specified peptide as well as
; OTHER INFORMATION: peptides including the specified peptide as a fragment there
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "cysteine residues at positions 2
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"
; US-08-381-984-25

Query Match 97.0%; Score 65; DB 1; Length 20;
Best Local Similarity 91.7%; Pred. No. 4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
DB 2 CFQWRNRKVR 13

RESULT 12
US-09-508-734-4
; Sequence 4, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; TITLE OF INVENTION: useful microorganism thereof
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KE99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KRI998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 4

```

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match
Best Local Similarity 97.0%; Score 65; DB 4; Length 22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match
Best Local Similarity 97.0%; Score 65; DB 4; Length 24;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 3 CFQWRNMRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match
Best Local Similarity 97.0%; Score 65; DB 1; Length 25;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRNMRKVR 12
Db 4 CFQWRNMRKVR 15
```

## RESULT 15

US-07-891-174-10  
; Sequence 10, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 4  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 21"  
; FEATURE:  
; NAME/KEY: modified site  
; LOCATION: 21

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "thiol group of  
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
; OTHER INFORMATION: thiol group of Cys residue at location 4"  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
; US-07-891-174-10

Query Match 97.0%; Score 65; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXNNMKVR 12  
||| |||||  
Db 4 CFQWQXNNMKVR 15

Search completed: February 21, 2003, 08:04:25  
Job time : 8.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107B-88  
Perfect score: 67  
Sequence: 1 CFQWXRNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pdb.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pdb.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pdb.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	97.0	15	9	US-09-798-869-2
2	65	97.0	25	9	US-09-798-869-20
3	65	97.0	694	9	US-10-023-096-2
4	57	85.1	15	9	US-09-798-869-6
5	48	71.6	15	9	US-09-798-869-3
6	48	71.6	25	9	US-09-798-869-23
7	40	59.7	15	9	US-09-798-869-7
8	39	58.2	15	9	US-09-798-869-4
9	39	58.2	25	9	US-09-798-869-22
10	37	55.2	489	9	US-09-888-320-2
11	36	53.7	15	9	US-09-798-869-8
12	36	53.7	15	9	US-09-798-869-29
13	36	53.7	15	9	US-09-798-869-30
14	35	52.2	21	10	US-09-864-761-47985
15	34	50.7	86	9	US-09-738-626-5715
16	34	50.7	95	10	US-09-764-864-1031
17	34	50.7	338	9	US-09-978-295A-119
18	34	50.7	338	9	US-09-978-697-119
19	34	50.7	338	9	US-09-978-192A-119

20	34	50.7	338	9	US-09-999-832A-119
21	34	50.7	338	9	US-09-978-189-119
22	34	50.7	553	9	US-09-796-753-14
23	34	50.7	553	10	US-09-981-649A-6
24	34	50.7	553	10	US-09-981-649A-24
25	34	50.7	554	10	US-09-981-649A-30
26	34	50.7	554	10	US-09-981-649A-32
27	34	50.7	556	10	US-09-981-649A-28
28	34	50.7	559	10	US-09-981-649A-28
29	34	50.7	846	9	US-10-051-409-4
30	34	50.7	1212	9	US-10-219-248-3
31	34	50.7	1212	9	US-10-219-247-3
32	34	50.7	1212	10	US-09-855-722-3
33	34	50.7	1238	9	US-10-219-248-5
34	34	50.7	1238	9	US-10-219-247-5
35	34	50.7	1238	10	US-09-855-722-5
36	34	50.7	1238	10	US-09-944-849-4
37	33	49.3	40	10	US-09-864-761-46393
38	33	49.3	62	10	US-09-815-242-12129
39	33	49.3	62	10	US-09-815-242-13026
40	33	49.3	333	9	US-09-796-753-26
41	33	49.3	1258	10	US-09-867-852-107
42	33	49.3	26926	9	US-09-759-508B-2
43	32	47.8	28	10	US-09-864-761-36085
44	32	47.8	31	9	US-09-956-206A-13
45	32	47.8	344	9	US-09-735-056-33

## ALIGNMENTS

## RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 97.0%; Score 65; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 1.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
Db 3 CFQWXRNRKVR 14

## RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON

APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 97.0%; Score 65; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 2.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
Db 3 CFQWXRNRKVR 14

## RESULT 3

US-10-023-096-2  
Sequence 2, Application US/10023096  
Patent No. US20020160941A1  
GENERAL INFORMATION:  
APPLICANT: Kruzel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,096  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 97.0%; Score 65; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.00066;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
Db 22 CFQWXRNRKVR 33

## RESULT 4

US-09-798-869-6  
Sequence 6, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 85.1%; Score 57; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00044;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12  
Db 3 CFQWXRNRKVR 14

## RESULT 5

US-09-798-869-3  
Sequence 3, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15  
TYPE: PRT  
ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 71.6%; Score 48; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.016;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

# RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 71.6%; Score 48; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.026;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

# RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 59.7%; Score 40; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.4;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CYQWQWRMRKL 13

# RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 58.2%; Score 39; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.6;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQWRMRKL 13

# RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 58.2%; Score 39; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 0.97;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
|:|:|:|:|:|:  
Db 3 CLRQWQWRMRKL 13



RESULT 10  
 US-09-888-320-2  
 ; Sequence 2, Application US/09888320  
 ; Publication No. US20030013090A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barry III, Clifton E.  
 ; APPLICANT: DeBarber, Andrea E.  
 ; APPLICANT: Mlului, Khisimuizi  
 ; APPLICANT: Bekker, Linda-Gail  
 ; APPLICANT: The Government of the United States of America  
 ; APPLICANT: as represented by The Secretary of the  
 ; APPLICANT: Department of Health and Human Services  
 ; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
 ; FILE REFERENCE: 015280-413100US  
 ; CURRENT APPLICATION NUMBER: US/09/888,320  
 ; CURRENT FILING DATE: 2001-06-22  
 ; PRIOR APPLICATION NUMBER: US 60/214,187  
 ; PRIOR FILING DATE: 2000-06-26  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 489  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: wild-type Etaa monoxygenase (Rv3854C, EthA)  
 US-09-888-320-2

Query Match 55.2%; Score 37; DB 9; Length 489;  
 Best Local Similarity 54.5%; Pred. No. 36;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
 :|||:  
 Db 253 CQKQWRMRKN 263

RESULT 11  
 US-09-798-869-8  
 ; Sequence 8, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
 ; OTHER INFORMATION: sequence)  
 US-09-798-869-8

Query Match 53.7%; Score 36; DB 9; Length 15;  
 Best Local Similarity 54.5%; Pred. No. 2;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
 :|||:  
 Db 3 CLRQWQWRMRKN 13

RESULT 12  
 US-09-798-869-29  
 ; Sequence 29, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 29  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: BOVINE  
 US-09-798-869-29

Query Match 53.7%; Score 36; DB 9; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 2;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
 :|||:  
 Db 3 CFRQWRMRKKL 13

RESULT 13  
 US-09-798-869-30  
 ; Sequence 30, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 30  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: BOVINE  
 US-09-798-869-30

Query Match 53.7%; Score 36; DB 9; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 2;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNRKV 11  
 :|||:  
 Db 3 CFRQWRMRKKL 13

RESULT 14  
 US-09-864-761-47985  
 ; Sequence 47985, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47985
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096701.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EST_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06
US-09-864-761-47985

```

```

Query Match      52.2%; Score 35; DB 10; Length 21;
Best Local Similarity 83.3%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CFQWXR 6
    |||||
Db 16 CFQWRR 21

```

```

RESULT 15
US-09-738-626-5715
; Sequence 5715, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:

```

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5715
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5715

```

```

Query Match      50.7%; Score 34; DB 9; Length 86;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 FQWXRNRKVR 12
    |::|::|::|
Db 73 FEYRRQLRKIR 83

```

```

Search completed: February 21, 2003, 08:11:55
Job time : 6.88372 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107B-88

Perfect score: 67

Sequence: 1 CFQWRNMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	97.0	711	1 TFHUL	lactotransferrin p
2	48	71.6	708	2 J2323	lactoferrin - goat
3	45	67.2	33	2 S52107	lactoferrin - sheep
4	41	61.2	511	2 A30858	hypothetical prote
5	39	58.2	275	2 T22597	hypothetical prote
6	39	58.2	707	1 A28438	lactoferrin precu
7	38	56.7	531	2 A84471	En/Spm-like transp
8	38	56.7	536	2 T24218	hypothetical prote
9	38	56.7	4568	2 T08030	apoptin beta heavy
10	37	55.2	274	2 B60950	apolipoprotein B-1
11	37	55.2	489	2 C70655	probable monooxyge
12	37	55.2	501	2 T39801	hypothetical sh3-c
13	37	55.2	584	2 C84325	hypothetical prote
14	36	53.7	124	2 C96582	F511.22 [imported
15	36	53.7	298	2 A23466	hypothetical prote
16	36	53.7	361	2 A24470	hypothetical prote
17	36	53.7	365	2 T37477	probable proteinas
18	36	53.7	369	2 B36418	MHC class I histoc
19	36	53.7	428	2 F91660	MAP1 protein - myx
20	36	53.7	742	2 T25415	histidyl-trna synt
21	36	53.7	749	2 A45687	hypothetical prote
22	36	53.7	2700	2 D88450	outer capsid prote
23	35	52.2	114	2 D33876	protein F21H1.2 [
24	35	52.2	205	2 E90094	carcinoembryonic a
25	35	52.2	206	2 H97451	26S proteasome SU
26	35	52.2	206	2 AB2670	pyridoxamine 5'-ph
27	35	52.2	208	2 AG3441	pyridoxamine 5'-ph
28	35	52.2	376	2 S67085	probable pyridoxam
29	35	52.2	649	2 AB2154	hypothetical prote

30	35	52.2	932	2 T28820	hypothetical prote
31	35	52.2	966	1 P1BYBB	RNA 1a protein - b
32	35	52.2	1135	2 T14803	phytochrome C - so
33	35	52.2	1174	2 C97686	pyruvate carboxyla
34	35	52.2	1174	2 AE2911	pyruvate carboxyla
35	35	52.2	1213	2 T41378	probable helicase
36	35	52.2	1432	2 B85431	trichohyalin like
37	35	52.2	1804	2 T30563	resistance protein
38	35	52.2	6642	2 T29757	protein UNC-89 - C
39	34	50.7	214	2 S07989	vif protein - simi
40	34	50.7	224	2 D89836	hypothetical prote
41	34	50.7	249	2 A13401	arginyltransferase
42	34	50.7	255	2 E87515	signal peptidase I
43	34	50.7	289	2 G86403	33.3K hypothetical
44	34	50.7	323	2 C82234	conserved hypothet
45	34	50.7	335	2 T33211	hypothetical prote

## ALIGNMENTS

### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74:

R:Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AA60324.1; PID:G467237

R:Rev. M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148 'T', 150-422 'C' 424-711 <REY>

A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A:Title: Differential molecular mechanism of the estrogen action that regulates lactoferrin

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:G263311; PIDN:AA624877.1; PID:G263312

A:Experimental source: placenta

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A:Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <ST1>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-28, 'X', 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A:Reference number: S07160; PMID:88001031; PMID:3477300  
 A:Accession: S07160  
 A:Molecule type: mRNA  
 A:Residues: 436-487, 'A', 489-711 <RAD>  
 A:Cross-references: EMBL:M18642; NID:G186815; PID:AAA8665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A:Reference number: A61169; PMID:91235214; PMID:1674448  
 A:Accession: A61169  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 3-701, 'SWKPVN' <PAN>  
 A:Experimental source: Jolles breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1984  
 A>Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A:Reference number: A31000; PMID:85076667; PMID:6510420  
 A:Accession: A31000  
 A:Molecule type: protein  
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-400  
 A>Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L. Eur. J. Biochem. 241, 303-308, 1996  
 A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity chromatography  
 A:Reference number: S74119; PMID:97054624; PMID:8898921  
 A:Accession: S74119  
 A:Molecule type: protein  
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A:Experimental source: neutrophil granulocytes  
 C:Genetics:  
 A:Gene: GDB:ITF  
 A:Cross-references: GDB:119368; OMIM:150210  
 A:Map position: 3q21-3q23  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-711/Product: lactotransferrin #status experimental <MAT>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status experimental  
 F:157, 458/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:368-400, 378-391, 423-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #status experimental

Query Match 97.0%; Score 65; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.0005;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 12  
 |||||  
 Db 39 CFQWQRNRKVR 50

RESULT 2  
 JC2323  
 lactoferrin - goat  
 C:Species: Capra aegagrus hircus (domestic goat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C:Accession: JC2323  
 R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P. Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A:Reference number: JC2323; PMID:94380047; PMID:8093048  
 A:Accession: JC2323  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <LEP>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:359-696/Domain: transferrin repeat homology <TRH2>  
 F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.6%; Score 48; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.65;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 11  
 |||||  
 Db 38 CFQWQRNRKVL 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C:Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Magliore-Sancour, D.; Fiat, A.M. Biochim. Biophys. Acta 1243, 25-32, 1995  
 A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet aggregation  
 A:Reference number: S52107; PMID:95127729; PMID:7827104  
 A:Accession: S52107  
 A>Status: preliminary  
 A:Molecule type: Protein  
 A:Residues: 1-33 <QIA>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication

Query Match 67.2%; Score 45; DB 2; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.12;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 11  
 |||||  
 Db 19 CFQWQKQKVRKL 29

RESULT 4  
 AB0858  
 hypothetical protein STY3070 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 C>Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AB0858  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; et al.  
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A:Reference number: AB0858; PMID:11677608  
 A:Accession: AB0858  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-511 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:G16504016; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY3070

Query Match 61.2%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 9.1;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFQWXRNRKVR 12  
 |||||  
 Db 350 CFQWDMNKAKVR 361

RESULT 5  
 T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T22597

R;Dobson, R.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19587  
A;Accession: T22597  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-275 <WIL>  
A;Cross-references: EMBL:Z81089; PIDN:CA803137.1; GSPDB:GN00028; CESP:F53H4.4  
A;Experimental source: clone F53H4  
C;Genetics:  
A;Gene: CESP:F53H4.4  
A;Map position: X  
A;Introns: 67/1; 153/1  
C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 58.2%; Score 39; DB 2; Length 275;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWXRNMKVR 12  
|||:||||  
Db 262 FQWXRNMKVR 272

RESULT 6  
A28438  
lactoferrin precursor - mouse  
N;Alternate names: lactotransferrin  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A28438, A41205  
R;Pentecost, B.T.; Teng, C.T.  
J. Biol. Chem. 262, 10134-10139, 1987  
A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory gland  
A;Reference number: A92596; MUID:87280033; PMID:3611056  
A;Accession: A28438  
A;Molecule type: mRNA  
A;Residues: 3-707 <PEN>  
A;Cross-references: EMBL:J03298  
R;Liu, Y.; Teng, C.T.  
J. Biol. Chem. 266, 21880-21885, 1991  
A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
A;Reference number: A41205; MUID:92042099; PMID:1939212  
A;Accession: A41205  
A;Molecule type: DNA  
A;Residues: 1-15 <LIU>  
A;Cross-references: GB:M74778  
C;Superfamily: transferrin; transferrin repeat homology  
C;Keywords: duplication; glycoprotein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-707/Product: lactotransferrin #status predicted <MAT>  
F;358-695/Domain: transferrin repeat homology <TRH2>  
F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.2%; Score 39; DB 1; Length 707;  
Best Local Similarity 54.5%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRNMKVR 11  
|||:||||  
Db 37 CLRQWQMKVR 47

RESULT 7  
A94471  
En/Spm-like transposon protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
A;Accession: A94471  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

Query Match 56.7%; Score 38; DB 2; Length 536;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRNMKVR 12  
|||:||||  
Db 293 CIDWGRDRKVK 304

RESULT 9  
T08030  
dynein beta heavy chain - Chlamydomonas reinhardtii  
C;Species: Chlamydomonas reinhardtii  
C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
A;Accession: T08030  
R;Mitchell, D.R.; Brown, K.S.  
J. Cell Sci. 107, 635-644, 1994  
A;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
A;Reference number: Z16302; MUID:94274778; PMID:8006077  
A;Accession: T08030  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-4568 <MIT>  
A;Cross-references: EMBL:U02963; NID:9409965; PIDN:AAA19956.1; PID:G514215  
A;Experimental source: strain 21gr  
C;Genetics:  
A;Gene: ODA4  
A;Map position: IX  
A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3; 3334/3; 3686/3; 3882/3; 4240/3  
C;Superfamily: dynein heavy chain, ciliary  
C;Keywords: nucleotide binding; P-loop  
F;1919-1926/Region: nucleotide-binding motif A (P-loop)

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: A84471  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-531 <STO>  
A;Cross-references: GB:AE002093; NID:94586022; PIDN:AAD25641.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g05650  
A;Map position: 2

Query Match 56.7%; Score 38; DB 2; Length 531;  
Best Local Similarity 75.0%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWXRNMK 10  
|||:||||  
Db 501 QWXRNMK 508

RESULT 8  
T24218  
hypothetical protein R13G10.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T24218  
R;Gardner, A.  
submitted to the EMBL Data Library, August 1994  
A;Reference number: Z19857  
A;Accession: T24218  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-536 <WIL>  
A;Cross-references: EMBL:Z35602; PIDN:CAA84671.1; GSPDB:GN00021; CESP:R13G10.2  
A;Experimental source: clone R13G10  
C;Genetics:  
A;Gene: CESP:R13G10.2  
A;Map position: 3  
A;Introns: 64/3; 194/1; 404/3

Query Match 56.7%; Score 38; DB 2; Length 536;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRNMKVR 12  
|||:||||  
Db 293 CIDWGRDRKVK 304

RESULT 9  
T08030  
dynein beta heavy chain - Chlamydomonas reinhardtii  
C;Species: Chlamydomonas reinhardtii  
C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
A;Accession: T08030  
R;Mitchell, D.R.; Brown, K.S.  
J. Cell Sci. 107, 635-644, 1994  
A;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
A;Reference number: Z16302; MUID:94274778; PMID:8006077  
A;Accession: T08030  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-4568 <MIT>  
A;Cross-references: EMBL:U02963; NID:9409965; PIDN:AAA19956.1; PID:G514215  
A;Experimental source: strain 21gr  
C;Genetics:  
A;Gene: ODA4  
A;Map position: IX  
A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3; 3334/3; 3686/3; 3882/3; 4240/3  
C;Superfamily: dynein heavy chain, ciliary  
C;Keywords: nucleotide binding; P-loop  
F;1919-1926/Region: nucleotide-binding motif A (P-loop)